Using Docker with Cisco NX-OS

About Docker with Cisco NX-OS

Docker provides a way to run applications securely isolated in a container, packaged with all its dependencies and libraries. See https://docs.docker.com/ for more information on Docker.

Beginning with Cisco NX-OS Release 9.2(1), support is now added for using Docker within Cisco NX-OS on a switch.

The version of Docker that is included on the switch is CE 18.09.0. The Docker daemon is not running by default. You must start it manually or set it up to automatically restart when the switch boots up.

This section describes how to enable and use Docker in the specific context of the switch environment. Refer to the Docker documentation at https://docs.docker.com/ for details on general Docker usage and functionality.

Guidelines and Limitations for Docker

Following are the guidelines and limitations for using Docker on Cisco NX-OS on a switch:

- If you are running a third-party DHCPD server in Docker, there might be issues with offers reaching the client if used along with SVI. A possible workaround is to use broadcast responses.
- Docker functionality is supported on the Cisco Nexus 9000Cisco Nexus 3000 Series switches with at least 8 GB of system RAM.

Prerequisites for Setting Up Docker Containers Within Cisco NX-OS

Following are the prerequisites for using Docker on Cisco NX-OS on a switch:

- Enable the host Bash shell. To use Docker on Cisco NX-OS on a switch, you must be the root user on the host Bash shell:

```
switch# configure terminal
    Enter configuration commands, one per line. End with CNTL/Z.
    switch(config)# feature bash-shell
```

- If the switch is in a network that uses an HTTP proxy server, the `http_proxy` and `https_proxy` environment variables must be set up in `/etc/sysconfig/docker`. For example:

```
export http_proxy=http://proxy.esl.cisco.com:8080
export https_proxy=http://proxy.esl.cisco.com:8080
```

- Verify that the switch clock is set correctly, or you might see the following error message:

```
x509: certificate has expired or is not yet valid
```

- Verify that the domain name and name servers are configured appropriately for the network and that it is reflected in the`/etc/resolv.conf` file:

```
switch# conf t
     Enter configuration commands, one per line. End with CNTL/Z.
     switch(config)# vrf context management
     switch(config-vrf)# ip domain-name ?
     WORD Enter the default domain (Max Size 64)

     switch(config-vrf)# ip name-server ?
     A.B.C.D Enter an IPv4 address
     A:B::C:D Enter an IPv6 address

root@switch# cat /etc/resolv.conf
domain cisco.com #bleed
nameserver 171.70.168.183 #bleed
root@switch#
```

Starting the Docker Daemon

When you start the Docker daemon for the first time, a fixed-size backend storage space is carved out in a file called `dockerpart` on the bootflash, which is then mounted to `/var/lib/docker`. If necessary, you can adjust the default size of this space by editing `/etc/sysconfig/docker` before you start the Docker daemon for the first time. You can also resize this storage space if necessary as described later on.

To start the Docker daemon:

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Start the Docker daemon.

```
root@switch# service docker start
```

Step 3 Check the status.

```
root@switch# service docker status
dockerd (pid 3597) is running...
root@switch#
```

Note Once you start the Docker daemon, do not delete or tamper with the `dockerpart` file on the bootflash since it is critical to the docker functionality.

```
switch# dir bootflash:dockerpart
2000000000 Mar 14 12:50:14 2018 dockerpart
```

Configure Docker to Start Automatically

You can configure the Docker daemon to always start up automatically when the switch boots up.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Use the `chkconfig` utility to make the Docker service persistent.

```
root@switch# chkconfig --add docker
root@n9k-2#
```

Step 3 Use the `chkconfig` utility to check the Docker service settings.

```
root@switch# chkconfig --list | grep docker
docker 0:off 1:off 2:on 3:on 4:on 5:on 6:off
root@switch#
```

Step 4 To remove the configuration so that Docker does not start up automatically:

```
root@switch# chkconfig --del docker
root@switch# chkconfig --list | grep docker
```

```
root@switch#
```

Starting Docker Containers: Host Networking Model

If you want Docker containers to have access to all the host network interfaces, including data port and management, start the Docker containers with the `--network host` option. The user in the container can switch between the different network namespaces at `/var/run/netns` (corresponding to different VRFs configured in Cisco NX-OS) using the `ip netns exec <net_namespace> <cmd>`.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Start the Docker container.

Following is an example of starting an Alpine Docker container on the switch and viewing all the network interfaces. The container is launched into the management network namespace by default.

```
root@switch# docker run --name=alpinerun -v /var/run/netns:/var/run/netns:ro,rslave --rm --network
host --cap-add SYS_ADMIN -it alpine
/ # apk --update add iproute2
fetch http://dl-cdn.alpinelinux.org/alpine/v3.7/main/x86_64/APKINDEX.tar.gz
fetch http://dl-cdn.alpinelinux.org/alpine/v3.7/community/x86_64/APKINDEX.tar.gz
(1/6) Installing libelf (0.8.13-r3)
(2/6) Installing libmnl (1.0.4-r0)
(3/6) Installing jansson (2.10-r0)
(4/6) Installing libnftnl-libs (1.0.8-r1)
(5/6) Installing iptables (1.6.1-r1)
(6/6) Installing iproute2 (4.13.0-r0)
Executing iproute2-4.13.0-r0.post-install
Executing busybox-1.27.2-r7.trigger
OK: 7 MiB in 17 packages
/ #
/ # ip netns list
management
default
/ #
/ # ip address
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default
link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
inet 127.0.0.1/8 scope host lo
valid_lft forever preferred_lft forever
inet6 ::1/128 scope host
valid_lft forever preferred_lft forever
2: tunl0@NONE: <NOARP> mtu 1480 qdisc noop state DOWN group default
link/ipip 0.0.0.0 brd 0.0.0.0
3: gre0@NONE: <NOARP> mtu 1476 qdisc noop state DOWN group default
link/gre 0.0.0.0 brd 0.0.0.0
...
/ #
/ # ip netns exec default ip address
```

```
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default
link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
inet 127.0.0.1/16 scope host lo
valid_lft forever preferred_lft forever
2: dummy0: <BROADCAST,NOARP> mtu 1500 qdisc noop state DOWN group default
link/ether 42:0d:9b:3c:d4:62 brd ff:ff:ff:ff:ff:ff
3: tunl0@NONE: <NOARP> mtu 1480 qdisc noop state DOWN group default
link/ipip 0.0.0.0 brd 0.0.0.0
...
```

Starting Docker Containers: Bridged Networking Model

If you want Docker containers to only have external network connectivity (typically through the management interface) and you don't necessarily care about visibility into a specific data port or other switch interface, you can start the Docker container with the default Docker bridged networking model. This is more secure than the host networking model described in the previous section since it also provides network namespace isolation.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Start the Docker container.

Following is an example of starting an Alpine Docker container on the switch and installing the `iproute2` package.

```
root@switch# docker run -it --rm alpine
/ # apk --update add iproute2
fetch http://dl-cdn.alpinelinux.org/alpine/v3.7/main/x86_64/APKINDEX.tar.gz
fetch http://dl-cdn.alpinelinux.org/alpine/v3.7/community/x86_64/APKINDEX.tar.gz
(1/6) Installing libelf (0.8.13-r3)
(2/6) Installing libmnl (1.0.4-r0)
(3/6) Installing jansson (2.10-r0)
(4/6) Installing libnftnl-libs (1.0.8-r1)
(5/6) Installing iptables (1.6.1-r1)
(6/6) Installing iproute2 (4.13.0-r0)
Executing iproute2-4.13.0-r0.post-install
Executing busybox-1.27.2-r7.trigger
OK: 7 MiB in 17 packages
/ #
/ # ip netns list
/ #
```

Step 3 Determine if you want to set up user namespace isolation.

For containers using the bridged networking model, you can also set up user namespace isolation to further improve security. See Securing Docker Containers With User namespace Isolation, on page 11 for more information.

You can use standard Docker port options to expose a service from within the container, such as sshd. For example:

```
root@switch# docker run -d -p 18877:22 --name sshd_container sshd_ubuntu
```

This maps port 22 from within the container to port 18877 on the switch. The service can now be accessed externally through port 18877, as shown in the following example:

```
root@ubuntu-vm# ssh root@ip_address -p 18887
```

Mounting the bootflash and volatile Partitions in the Docker Container

You can make the `bootflash` and `volatile` partitions visible in the Docker container by passing in the `-v /bootflash:/bootflash` and `-v /volatile:/volatile` options in the run command for the Docker container. This is useful if the application in the container needs access to files shared with the host, such as copying a new NX-OS system image to bootflash.

Note This **-v** command option allows for any directory to be mounted into the container and may result in information leaking or other accesses that may impact the operation of the NX-OS system. Limit this to resources such as `/bootflash` and `/volatile` that are already accessible using NX-OS CLI.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Pass in the `-v /bootflash:/bootflash` and `-v /volatile:/volatile` options in the run command for the Docker container.

```
root@switch# docker run -v /bootflash:/bootflash -v /volatile:/volatile -it --rm alpine
/# ls /
bin         etc         media       root        srv         usr
bootflash   home        mnt         run         sys         var
dev         lib         proc        sbin        tmp         volatile
/ #
```

Enabling Docker Daemon Persistence on Enhanced ISSU Switchover

You can have both the Docker daemon and any running containers persist on an Enhanced ISSU switchover. This is possible since the bootflash on which the backend Docker storage resides is the same and shared between both Active and Standby supervisors.

The Docker containers are disrupted (restarted) during the switchover, so they will not be running continuously.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Before starting the switchover, use the `chkconfig` utility to make the Docker service persistent.

```
root@switch# chkconfig --add docker
root@n9k-2#
```

Step 3 Start any containers using the `--restart unless-stopped` option so that they will be restarted automatically after the switchover.

The following example starts an Alpine container and configures it to always restart unless it is explicitly stopped or Docker is restarted:

```
root@switch# docker run -dit --restart unless-stopped alpine
root@n9k-2#
```

The Docker containers are disrupted (restarted) during the switchover, so they will not be running continuously.

Enabling Docker Daemon Persistence on the Cisco Nexus Platform Switches Switchover

You can have both the Docker daemon and any running containers persist on a switchover between two separate physical supervisors with distinct bootflash partitions. However, for the Cisco Nexus switches, the bootflash partitions on both supervisors are physically separate. You will therefore need to copy the `dockerpart` file manually to the standby supervisor before performing the switchover.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Start any containers using the `--restart unless-stopped` option so that they will be restarted automatically after the switchover.

The following example starts an Alpine container and configures it to always restart unless it is explicitly stopped or Docker is restarted:

```
root@switch# docker run -dit --restart unless-stopped alpine
root@n9k-2#
```

Note that the Docker containers will be disrupted (restarted) during the switchover, so they will not be running continuously.

Step 3 Before starting the switchover, use the `chkconfig` utility to make the Docker service persistent.

```
root@switch# chkconfig --add docker
root@n9k-2#
```

Step 4 Copy the Docker backend storage partition from the active to the standby supervisor bootflash:

```
root@switch# service docker stop
Stopping dockerd: dockerd shutdown

root@switch# cp /bootflash/dockerpart /bootflash_sup-remote/

root@switch# service docker start
```

Resizing the Docker Storage Backend

After starting or using the Docker daemon, you can grow the size of the Docker backend storage space according to your needs.

Procedure

Step 1 Disable the Guest Shell.

If you do not disable the Guest Shell, it may interfere with the resize.

```
switch# guestshell disable
You will not be able to access your guest shell if it is disabled. Are you sure you want to disable
 the guest shell? (y/n) [n] y
switch# 2018 Mar 15 17:16:55 switch %$ VDC-1 %$ %VMAN-2-ACTIVATION_STATE: Deactivating virtual
service 'guestshell+'
2018 Mar 15 17:16:57 switch %$ VDC-1 %$ %VMAN-2-ACTIVATION_STATE: Successfully deactivated virtual
 service 'guestshell+'
```

Step 2 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 3 Get information on the current amount of storage space available.

```
root@switch# df -kh /var/lib/docker
Filesystem Size Used Avail Use% Mounted on
/dev/loop12 1.9G 7.6M 1.8G 1% /var/lib/docker
root@n9k-2#
```

Step 4 Stop the Docker daemon.

```
root@switch# service docker stop
Stopping dockerd: dockerd shutdown
```

Step 5 Get information on the current size of the Docker backend storage space (`/bootflash/dockerpart`).

```
root@switch# ls -l /bootflash/dockerpart
-rw-r--r-- 1 root root 2000000000 Mar 15 16:53 /bootflash/dockerpart
root@n9k-2#
```

Step 6 Resize the Docker backend storage space.

For example, the following command increases the size by 500 megabytes:

```
root@switch# truncate -s +500MB /bootflash/dockerpart
root@n9k-2#
```

Step 7 Get updated information on the size of the Docker backend storage space to verify that the resizing process was completed successfully.

For example, the following output confirms that the size of the Docker backend storage was successfully increased by 500 megabytes:

```
root@switch# ls -l /bootflash/dockerpart
-rw-r--r-- 1 root root 2500000000 Mar 15 16:54 /bootflash/dockerpart
root@n9k-2#
```

Step 8 Check the size of the filesystem on `/bootflash/dockerpart`.

```
root@switch# e2fsck -f /bootflash/dockerpart
e2fsck 1.42.9 (28-Dec-2013)
Pass 1: Checking inodes, blocks, and sizes
Pass 2: Checking directory structure
Pass 3: Checking directory connectivity
Pass 4: Checking reference counts
Pass 5: Checking group summary information
/bootflash/dockerpart: 528/122160 files (0.6% non-contiguous), 17794/488281 blocks
```

Step 9 Resize the filesystem on `/bootflash/dockerpart`.

```
root@switch# /sbin/resize2fs /bootflash/dockerpart
resize2fs 1.42.9 (28-Dec-2013)
Resizing the filesystem on /bootflash/dockerpart to 610351 (4k) blocks.
The filesystem on /bootflash/dockerpart is now 610351 blocks long.
```

Step 10 Check the size of the filesystem on `/bootflash/dockerpart` again to confirm that the filesystem was successfully resized.

```
root@switch# e2fsck -f /bootflash/dockerpart
e2fsck 1.42.9 (28-Dec-2013)
Pass 1: Checking inodes, blocks, and sizes
Pass 2: Checking directory structure
Pass 3: Checking directory connectivity
Pass 4: Checking reference counts
Pass 5: Checking group summary information
/bootflash/dockerpart: 528/154736 files (0.6% non-contiguous), 19838/610351 blocks
```

Step 11 Start the Docker daemon again.

```
root@switch# service docker start
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...
done.
Starting dockerd with args '--debug=true':
```

Step 12 Verify the new amount of storage space available.

```
root@switch# df -kh /var/lib/docker
Filesystem Size Used Avail Use% Mounted on
/dev/loop12 2.3G 7.6M 2.3G 1% /var/lib/docker
```

Step 13 Exit out of Bash shell.

```
root@switch# exit
logout
switch#
```

Step 14 Enable the Guest Shell, if necessary.

```
switch# guestshell enable

switch# 2018 Mar 15 17:12:53 switch %$ VDC-1 %$ %VMAN-2-ACTIVATION_STATE: Activating virtual service
 'guestshell+'
switch# 2018 Mar 15 17:13:18 switch %$ VDC-1 %$ %VMAN-2-ACTIVATION_STATE: Successfully activated
virtual service 'guestshell+'
```

Stopping the Docker Daemon

If you no longer wish to use Docker, follow the procedures in this topic to stop the Docker daemon.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Stop the Docker daemon.

```
root@switch# service docker stop
Stopping dockerd: dockerd shutdown
```

Step 3 Verify that the Docker daemon is stopped.

```
root@switch# service docker status
dockerd is stopped
root@switch#
```

Note You can also delete the `dockerpart` file on the bootflash at this point, if necessary:

```
switch# delete bootflash:dockerpart
Do you want to delete "/dockerpart" ? (yes/no/abort) y
switch#
```

Docker Container Security

Following are the Docker container security recommendations:

- Run in a separate user `namespace` if possible.
- Run in a separate network `namespace` if possible.
- Use cgroups to limit resources. An existing cgroup (`ext_ser`) is created to limit hosted applications to what the platform team has deemed reasonable for extra software running on the switch. Docker allows use of this and limiting per-container resources.
- Do not add unnecessary POSIX capabilities.

Securing Docker Containers With User namespace Isolation

For containers using the bridged networking model, you can also set up user namespace isolation to further improve security. See https://docs.docker.com/engine/security/userns-remap/ for more information.

Procedure

Step 1 Determine if a `dockremap` group already exists on your system.

A `dockremap` user must already be set up on your system by default. If the `dockremap` group doesn't already exist, follow these steps to create it.

a) Enter the following command to create the `dockremap` group:

```
root@switch# groupadd dockremap -r
```

b) Create the `dockremap` user, unless it already exists:

```
root@switch# useradd dockremap -r -g dockremap
```

c) Verify that the `dockremap` group and the `dockremap` user were created successfully:

```
root@switch# id dockremap
uid=999(dockremap) gid=498(dockremap) groups=498(dockremap)
root@switch#
```

Step 2 Add the desired re-mapped ID and range to the `/etc/subuid` and `/etc/subgid`.

For example:

```
root@switch# echo "dockremap:123000:65536" >> /etc/subuid
root@switch# echo "dockremap:123000:65536" >> /etc/subgid
```

Step 3 Using a text editor, add the `--userns-remap=default` option to the `other_args` field in the `/etc/sysconfig/docker` file.

For example:

```
other_args="-debug=true --userns-remap=default"
```

Step 4 Restart the Docker daemon, or start it if it is not already running, using `service docker [re]start`.

For example:

```
root@switch# service docker [re]start
```

Refer to the Docker documentation at https://docs.docker.com/engine/security/userns-remap/ for more information on configuring and using containers with user namespace isolation.

Moving the `cgroup` Partition

The `cgroup` partition for third-party services is `ext_ser`, which limits CPU usage to 25% per core. Cisco recommends that you run your Docker container under this `ext_ser` partition.

If the Docker container is run without the `--cgroup-parent=/ext_ser/` option, it can get up to the full 100% host CPU access, which can interfere with the regular operation of Cisco NX-OS.

Procedure

Step 1 Load Bash and become superuser.

```
switch# run bash sudo su -
```

Step 2 Run the Docker container under the `ext_ser` partition.

For example:

```
root@switch# docker run --name=alpinerun -v /var/run/netns:/var/run/netns:ro,rslave --rm --network
host --cgroup-parent=/ext_ser/ --cap-add SYS_ADMIN -it alpine
/ #
```

Adding Nodes to a Kubernetes Cluster

This topic describes how to add nodes to a Kubernetes cluster. In this example:

- The Kubernetes (Ubuntu) primary has an IP address of 10.122.197.246
- The switch software running as Docker containers has an IP address of 10.122.84.24

Note In the following examples, long single lines of text are broken up with the \ character to improve readability.

Procedure

Step 1 Run the following commands (on? for?) the Kubernetes (Ubuntu) primary.

a) Enter this command:

```
root@switch# docker run -d --net=host gcr.io/google_containers/etcd:2.2.1 /usr/local/bin/etcd
--listen-client-urls=http://0.0.0.0:4001 --advertise-client-urls=http://0.0.0.0:4001
--data-dir=/var/etcd/data
```

b) Enter this command:

```
root@switch# docker run -d --name=api --net=host --pid=host --privileged=true
gcr.io/google_containers/hyperkube:v1.2.2 /hyperkube apiserver --insecure-bind-address=0.0.0.0
--allow-privileged=true --service-cluster-ip-range=10.0.0.1/24 --etcd_servers=http://127.0.0.1:4001
 --v=2
```

c) Enter this command:

```
root@switch# docker run -d --name=kubs --volume=/:/rootfs:ro --volume=/sys:/sys:ro
--volume=/dev:/dev --volume=/var/lib/docker/:/var/lib/docker:rw
--volume=/var/lib/kubelet/:/var/lib/kubelet:rw --volume=/var/run:/var/run:rw --net=host --pid=host
 --privileged=true gcr.io/google_containers/hyperkube:v1.2.2 /hyperkube kubelet
--allow-privileged=true --hostname-override="127.0.0.1" --address="0.0.0.0"
--api-servers=http://0.0.0.0:8080 --cluster_dns=10.0.0.10 --cluster_domain=cluster.local
--config=/etc/kubernetes/manifests-multi
```

d) Enter this command:

```
root@switch# docker run -d --name=proxy --net=host --privileged
gcr.io/google_containers/hyperkube:v1.2.2 /hyperkube proxy --master=http://0.0.0.0:8080 --v=2
```

e) Enter this command:

```
root@switch# export KUBERNETES_MASTER=http://10.122.197.246:8080
```

f) Enter this command:

```
root@switch# curl -o /usr/bin/kubectl
http://storage.googleapis.com/kubernetes-release/release/v1.2.2/bin/linux/amd64/kubectl
```

g) Enter this command:

```
root@switch# kubectl -s $KUBERNETES_MASTER create -f kube-system.json
```

h) Enter this command:

```
root@switch# kubectl -s $KUBERNETES_MASTER create -f skydns-rc.yaml
```

i) Enter this command:

```
root@switch# kubectl -s $KUBERNETES_MASTER create -f skydns-svc.yaml
```

j) Enter this command:

```
root@switch# kubectl -s $KUBERNETES_MASTER create –f dashboard.yaml
kubectl -s $KUBERNETES_MASTER cluster-info
```

k) Enter this command:

```
root@switch# kubectl -s $KUBERNETES_MASTER cluster-info
```

Step 2 Run the following steps (on? for?) the switch.

a) Enter this command:

```
root@switch# docker run -d --name=kubs --net=host --pid=host --privileged=true --volume=/:/rootfs:ro
 --volume=/sys:/sys:ro --volume=/dev:/dev --volume=/var/lib/docker/:/var/lib/docker:rw
--volume=/var/lib/kubelet/:/var/lib/kubelet:rw --volume=/var/run:/var/run:rw
gcr.io/google_containers/hyperkube:v1.2.2 /hyperkube kubelet --allow-privileged=true --containerized
 --enable-server --cluster_dns=10.0.0.10 --cluster_domain=cluster.local
--config=/etc/kubernetes/manifests-multi --hostname-override="10.122.84.34" --address=0.0.0.0
--api-servers=http://10.122.197.246:8080
```

b) Enter this command:

```
root@switch# docker run -d --name=proxy --net=host --privileged=true
gcr.io/google_containers/hyperkube:v1.2.2 /hyperkube proxy --master=http://10.122.197.246:8080
--v=2
```

Step 3 Run the following commands (on? for?) the Kubernetes (Ubuntu) primary to deploy an nginx app in a replication controller object.

a) Enter this command:

```
lab@rnimbalk-ubuntu1:~$ kubectl get node
NAME             STATUS    AGE
10.122.84.34     Ready     16m
127.0.0.1        Ready     22m
```

b) Enter this command:

```
lab@rnimbalk-ubuntu1:~$ kubectl apply –f replication.yaml
replicationcontroller "nginx" created
```

c) Enter this command:

```
lab@rnimbalk-ubuntu1:~$ kubectl describe –f replication.yaml
Name:           nginx
Namespace:      default
Image(s):       nginx
Selector:       app=nginx
Labels:         app=nginx
Replicas:       3 current / 3 desired
Pods Status:    3 Running / 0 Waiting / 0 Succeeded / 0 Failed
No volumes.
Events:
  FirstSeen     LastSeen        Count   From                            SubobjectPath   Type
     Reason                  Message
  ---------     --------        -----   ----                            -------------   --------
     ------                  -------
  17s           17s             1       {replication-controller }                       Normal
     SuccessfulCreate Created pod: nginx-zqfpz
```

```
17s           17s             1       {replication-controller }                       Normal
   SuccessfulCreate Created pod: nginx-jij40
17s           17s             1       {replication-controller }                       Normal
   SuccessfulCreate Created pod: nginx-loa0g
```

Docker Troubleshooting

These topics describe issues that can arise with Docker containers and provides possible resolutions.

Docker Fails to Start

Problem: Docker fails to start, showing an error message similar to the following:

```
switch# run bash
bash-4.3$ service docker start
Free bootflash: 39099 MB, total bootflash: 51771 MB
Carving docker bootflash storage: 2000 MB
2000+0 records in
2000+0 records out
2000000000 bytes (2.0 GB) copied, 22.3039 s, 89.7 MB/s
losetup: /dev/loop18: failed to set up loop device: Permission denied
mke2fs 1.42.9 (28-Dec-2013)
mkfs.ext4: Device size reported to be zero.  Invalid partition specified, or
        partition table wasn't reread after running fdisk, due to
        a modified partition being busy and in use.  You may need to reboot
        to re-read your partition table.

Failed to create docker volume
```

Possible Cause: You might be running Bash as an admin user instead of as a root user.

Solution: Determine if you are running Bash as an admin user instead of as a root user:

```
bash-4.3$ whoami
admin
```

Exit out of Bash and run Bash as root user:

```
bash-4.3$ exit
switch# run bash sudo su -
```

Docker Fails to Start Due to Insufficient Storage

Problem: Docker fails to start, showing an error message similar to the following, due to insufficient bootflash storage:

```
root@switch# service docker start
Free bootflash: 790 MB, total bootflash: 3471 MB
Need at least 2000 MB free bootflash space for docker storage
```

Possible Cause: You might not have enough free bootflash storage.

Solution: Free up space or adjust the *variable*_dockerstrg values in /etc/sysconfig/docker as needed, then restart the Docker daemon:

```
root@switch# cat /etc/sysconfig/docker
# Replace the below with your own docker storage backend boundary value (in MB)
# if desired.
boundary_dockerstrg=5000

# Replace the below with your own docker storage backend values (in MB) if
# desired. The smaller value applies to platforms with less than
# $boundary_dockerstrg total bootflash space, the larger value for more than
# $boundary_dockerstrg of total bootflash space.
small_dockerstrg=300
large_dockerstrg=2000
```

Failure to Pull Images from Docker Hub (509 Certificate Expiration Error Message)

Problem: The system fails to pull images from the Docker hub with an error message similar to the following:

```
root@switch# docker pull alpine
Using default tag: latest
Error response from daemon: Get https://registry-1.docker.io/v2/: x509: certificate has
expired or is not yet valid
```

Possible Cause: The system clock might not be set correctly.

Solution: Determine if the clock is set correctly or not:

```
root@n9k-2# sh clock
15:57:48.963 EST Thu Apr 25 2002
Time source is Hardware Calendar
```

Reset the clock, if necessary:

```
root@n9k-2# clock set hh:mm:ss { day month | month day } year
```

For example:

```
root@n9k-2# clock set 14:12:00 10 feb 2018
```

Failure to Pull Images from Docker Hub (Client Timeout Error Message)

Problem: The system fails to pull images from the Docker hub with an error message similar to the following:

```
root@switch# docker pull alpine
Using default tag: latest
Error response from daemon: Get https://registry-1.docker.io/v2/: net/http: request canceled
 while waiting for connection (Client.Timeout exceeded while awaiting headers)
```

Possible Cause: The proxies or DNS settings might not be set correctly.

Solution: Check the proxy settings and fix them, if necessary, then restart the Docker daemon:

```
root@switch# cat /etc/sysconfig/docker | grep proxy
#export http_proxy=http://proxy.esl.cisco.com:8080
```

```
#export https_proxy=http://proxy.esl.cisco.com:8080
root@switch# service docker [re]start
```

Check the DNS settings and fix them, if necessary, then restart the Docker daemon:

```
root@switch# cat /etc/resolv.conf
domain cisco.com #bleed
nameserver 171.70.168.183 #bleed
root@switch# # conf t
     Enter configuration commands, one per line. End with CNTL/Z.
     switch(config)# vrf context management
     switch(config-vrf)# ip domain-name ?
     WORD Enter the default domain (Max Size 64)

     switch(config-vrf)# ip name-server ?
     A.B.C.D Enter an IPv4 address
     A:B::C:D Enter an IPv6 address
root@switch# service docker [re]start
```

Docker Daemon or Containers Not Running On Switch Reload or Switchover

Problem: The Docker daemon or containers do not run after you have performed a switch reload or switchover.

Possible Cause: The Docker daemon might not be configured to persist on a switch reload or switchover.

Solution: Verify that the Docker daemon is configured to persist on a switch reload or switchover using the `chkconfig` command, then start the necessary Docker containers using the `--restart unless-stopped` option. For example, to start an Alpine container:

```
root@switch# chkconfig --add docker
root@switch#
root@switch# chkconfig --list | grep docker
docker 0:off 1:off 2:on 3:on 4:on 5:on 6:off
root@switch# docker run -dit --restart unless-stopped alpine
```

Resizing of Docker Storage Backend Fails

Problem: An attempt to resize the Docker backend storage failed.

Possible Cause: You might not have Guest Shell disabled.

Solution: Use the following command to determine if Guest Shell is disabled:

```
root@switch# losetup -a | grep dockerpart
root@n9k-2#
```

The command should not display any output if Guest Shell is disabled.

Enter the following command to disable the Guest Shell, if necessary:

```
switch# guestshell disable
```

If you still cannot resize the Docker backend storage, you can delete `/bootflash/dockerpart`, then adjust the `[small_]large_dockerstrg` in `/etc/sysconfig/docker`, then start Docker again to get a fresh Docker partition with the size that you want.

Docker Container Doesn't Receive Incoming Traffic On a Port

Problem: The Docker container doesn't receive incoming traffic on a port.

Possible Cause: The Docker container might be using a netstack port instead of a kstack port.

Solution: Verify that any ephemeral ports that are used by Docker containers are within the kstack range. Otherwise any incoming packets can get sent to netstack for servicing and dropped.

```
switch# show socket local-port-range
Kstack local port range (15001 - 58000)
Netstack local port range (58001 - 63535) and nat port range (63536 - 65535)
switch# conf t
Enter configuration commands, one per line. End with CNTL/Z.
switch(config)# sockets local-port-range <start_port> <end_port>
switch# run bash sudo su -
root@switch# cat /proc/sys/net/ipv4/ip_local_port_range
15001   58000
root@switch#
```

Unable to See Data Port And/Or Management Interfaces in Docker Container

Problem: You are unable to see the data port or management interfaces in the Docker container.

Solution:

- Verify that the Docker container is started in the host network namespace with all host namespaces mapped in using the `-v /var/run/netns:/var/run/netns:ro,rslave --network host` options.
- Once in the container, you will be in the management network namespace by default. You can use the `ip netns` utility to move to the default (`init`) network namespace, which has the data port interfaces. The `ip netns` utility might need to be installed in the container using dnf, apk, or something similar.

General Troubleshooting Tips

Problem: You have other issues with Docker containers that were not resolved using other troubleshooting processes.

Solution:

- Look for dockerd debug output in `/var/log/docker` for any clues as to what is wrong.
- Verify that your switch has 8 GB or more of RAM. Docker functionality is not supported on any switch that has less than 8 GB of RAM.