Docker Security

Docker

Host → Container

How to allow container to have the least possible permissions and access to the host in order to achieve isolation.

AppArmor

Running the image with an AppArmor profile enforced, we can isolate the host as much as we want for each task.

AppArmor profiles have rules that can restrict access of the running image to the host.

AppArmor

AppArmor profile:

- Match rule → allow
- No matching rule → deny

The more specific the profile is, the more strict it will be.

AppArmor

AppArmor profile:

1. Static Analysis
2. Dynamic Analysis

Static Analysis

A parser that extracts some first permissions on container out of

- Dockerfile
- .yml file (DockerCompose)

Static Analysis

Dockerfile

A recipe that tells Docker how to create the image for the container.

Dockerfile

- Mount: RUN mount dir & VOLUME /foo
- Network: EXPOSE ports
- Chmod: RUN chmod file
- Chown: RUN chown file

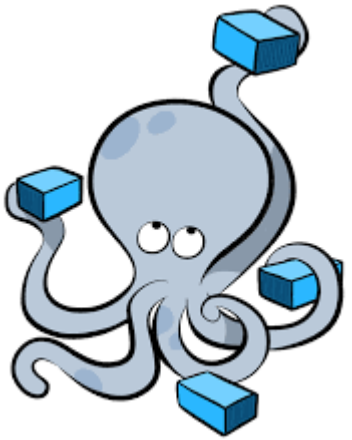
Dockerfile

- ~~Mount: RUN mount dir & VOLUME /foo~~
- ~~Network: EXPOSE ports~~
- Chmod: RUN chmod file
- Chown: RUN chown file

Mount is a privileged operation and for portability issues privileged operations are not allowed at the build phase, only at the runtime phase. For portability reasons again, we cannot expose certain ports at the build phase either.

DockerCompose

DockerCompose gives us parameters that should be given at runtime at docker, all included in a .yml file.



DockerCompose

- Network: ports
- Mount: volumes, volume_driver
- Capabilities: cap_add, cap_drop
- Devices: devices
- Resources: cgroup_parent, resources, ulimits (?)

AppArmor rules

Dockerfile

- Chmod rule
- Chown rule +

  Chown capability

DockerCompose

- Ports:

  Capability net_bind_service +

  network bind x to y
- Volumes:

  mount [<src>] [→ <mntpnt>]
- Capabilities add/drop:

  capability x / deny capability x
- Devices rule ???
- Resources, ulimits, cgroup_parent:

  rlimit <rlimit> <= <value>

AppArmor rules

1. Chmod & Chown rules:

Το Dockerfile φτιάχνει το image.

Μετά όμως, πρέπει να επιτρέψουμε στο container να έχει αυτό το permission?

Αν ναι, με την ίδια λογική, μήπως πρέπει να προσθέσουμε και για τις εντολές COPY και ADD access στο source με read και στο destination με write?

Αντίστοιχα και για WORKDIR path → path με write permission

AppArmor rules

2. Ports:

network [domain] [type] [protocol]

Για τα ports θεώρησα ότι χρειάζεται το **capability net_bind_service** (το είδα και σε έτοιμα profiles στα οποία το application είχε χρήση ports) και **network bind x to y** (δεν το έχω δει σε κάποιο profile)

AppArmor rules

3. Volumes:

Mount rule:

mount [<src>] [→ <mntpnt>]

Δεν νομίζω ότι χρειάζεται κάποιο capability.

Ίσως access στα paths των src και mntpnt?

(write access και στα δυο)

AppArmor rules

4. Capabilities:

- cap_add x: capability x
- cap_drop x: deny capability x

- ALL: Βάζω όλη τη λίστα των capabilities τα οποία υποστηρίζει τώρα το AppArmor.

  3 μόνο δεν υποστηρίζει σύμφωνα με το documentation: syslog, wake_alarm, audit_read τα οποία αγνοεί ο parser

AppArmor rules

5. Devices: match device host to device container

Δεν βρήκα κάποιο rule που να αφορά τα devices.

Σκέφτηκα να τα προσθέσω με mount rule, με paths τα paths των devices.

Είναι όμως η σχέση των device σαν mount?

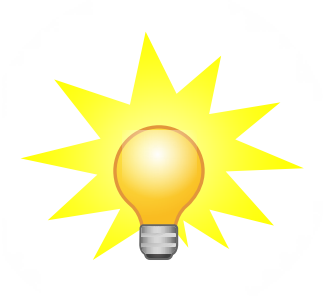
AppArmor rules

6. Resources: cgroup_parent, ulimits, resources

rlimit rule: rlimit <rlimit> <= <value>

Μας αφορά για το isolation?

Static Analysis



We encourage users to create a .yml Compose file even for a single service - task – so that the parser can extract rules before the dynamic analysis.

Dynamic Analysis

To enrich our profile we use AppArmor complain mode:

Profiling

- By hand
- By tools

with aa-complain

Dynamic Analysis

- Create a preliminary profile (static analysis) / Start aa-genprof in a terminal
- Put the profile into complain mode with aa-complain
- Execute test plan (start, stop, commands)
- Monitor the logs
- Adjust the profile
- Put the profile into enforce mode with aa-enforce

Dynamic Analysis

- Create a preliminary profile (static analysis) / Start aa-genprof in a terminal
- Put the profile into complain mode with aa-complain
- Execute test plan (start, stop, commands)
- Monitor the logs
- Adjust the profile
- Put the profile into enforce mode with aa-enforce

Τρέχουμε το container με ένα αντιπροσωπευτικό workload. Όπως???