# AN ANALYSIS OF SECURITY VULNERABILITIES IN CONTAINER IMAGES FOR SCIENTIFIC DATA ANALYSIS

Bhupinder Kaur

A thesis
in
The Department
of
Computer Science and Software Engineering (CSSE)

Presented in Partial Fulfillment of the Requirements
For the Degree of Master of Computer Science
Concordia University
Montréal, Québec, Canada

November 2020

# Concordia University
School of Graduate Studies

This is to certify that the thesis prepared

By: **Bhupinder Kaur**

Entitled: **An Analysis of Security Vulnerabilities in Container Images for Scientific Data Analysis**

and submitted in partial fulfillment of the requirements for the degree of

**Master of Computer Science**

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

_____ Chair

_____ Examiner

_____ Examiner

_____ Examiner

_____ Supervisor

Approved _____
Chair of Department or Graduate Program Director

_____ 20 _____ _____

Rama Bhat, Ph.D.,ing., FEIC, FCSME, FASME, Interim Dean
Faculty of Engineering and Computer Science

# Abstract

An Analysis of Security Vulnerabilities in Container Images for Scientific Data Analysis

Bhupinder Kaur

Software containers greatly facilitate the deployment and reproducibility of scientific data analyses on high-performance computing clusters (HPC). However, container images often contain outdated or unnecessary software packages, which increases the number of security vulnerabilities in the images and widens the attack surface of the infrastructure. This thesis presents a vulnerability analysis of container images for scientific data analysis. We compare results obtained with four vulnerability scanners, focusing on the use case of neuroscience data analysis, and quantifying the effect of image update and minification on the number of vulnerabilities. We find that container images used for neuroscience data analysis contain hundreds of vulnerabilities, that software updates remove about two thirds of these vulnerabilities, and that removing unused packages is also effective. We conclude with recommendations on how to build container images with a reduced amount of vulnerabilities.

# Acknowledgments

I would like to thank both my supervisors, Dr. Tristan Glatard and Dr. Aiman Hanna, for their unwavering guidance, support, valuable feedbacks, and patience throughout my thesis. I am glad and consider myself lucky that I have such experienced, motivated, and organized professors. I would also like to extend my gratitude to Mathieu Dugré for his help, efforts, and valuable inputs. I also had great pleasure of working with all my lab members, specially, Martin, Valérie, and Ali, you guys are so helpful and great colleagues. Many thanks to my family and friends for giving their emotional support and believing in me. Finally, big thanks to the Concordia University for providing me the chance to work with such supportive and knowledgeable supervisors, and giving me all the facilities and infrastructure to complete this work.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Containers are very lightweight, flexible, and more resource-efficient than virtual machines. They provide an operating system (OS) level virtualization, which removes the overhead of having an extra OS layer. However, containers tightly integrate with the host due to the sharing of the kernel, which raises security concerns. In this chapter, we provide a brief introduction of containers, mainly discussing two popular containerization techniques Docker and Singularity.

## 1.1 Containers

OS-level virtualization initially started in 1979 when Chroot was introduced. Chroot is a Unix system call, which changes the root directory of a process and its children processes. A chrooted process cannot access files outside the designated directory. As the time passed, this virtualization got evolved and took different forms of FreeBSD jails, Linux VServer project, and Solaris containers. However, in 2008, with the introduction of Linux containers (LXC) OS-level virtualization reached to an another level of popularity. LXC gained advantage over other Linux-based projects due to the integration of virtualization features into the upstream Linux kernel, which removed the need of applying patches to the Linux kernel and recompilation [2]. LXC relies on Linux kernel's control groups (cgroups) and namespaces. Cgroups control

the usage of system resources (memory, CPU, network, etc) among a group of processes. Namespaces define what a process can see by providing an isolated view of the operating system. For example, PID namespace isolate processes running in a container [3].

### 1.1.1 Docker

Docker was introduced in 2013 and has gained popularity since then. Docker is built on top of LXC. It automates application deployment, in addition to providing isolated environment. Docker allows software engineers to develop an application only once, on one system, and then use it on other systems without much effort. With the help of Docker, developers pack the application with all its dependencies, libraries, and tools, which can be run on any Linux server later. Docker containers share boot file-system and kernel with the host, which raises security concerns. Another issue is that Docker is not supported by traditional high-performance computing (HPC) resources. This is mainly due to the default configuration of Docker, which runs containers as root. Consequently, it becomes hard for system administrators to keep a record of which user is doing what [4, 5].

#### Docker Daemon

Docker engine is known as Docker daemon. It is a service that runs on the host operating system and is responsible for managing all Docker containers. It also allocates the host's resources to Docker containers. Like LXC, Docker daemon also uses *cgroups* and *namespaces* of Linux, and hence only support Linux operating system. Docker daemon requires root access to operate, and every container run by Docker is spawned as a child of root-owned Docker daemon. Therefore, users can gain escalated privileges by coercing Docker daemon, which creates serious security risks in shared environments. As a consequence, Docker is not supported in multi-user environments.

### Docker Images

A Docker image is a package which contains an application along with all required dependencies and libraries. All the steps of creating a particular image can be saved in a Docker file. Using Dockerfile, there are two ways to create a Docker image. The first is to use an existing Docker image as a base image and then build a new image from that one, by making changes in the filesystem, which is saved as a new layer of the Docker image. The second way is to build a new Docker image from scratch. Docker images are organized as a series of *layers* which are stacked on top of each other. Each layer consists of a filesystem diff that is introduced due to the changes made on the layer below it. All layers of the Docker image are identified by unique *layerIDs*. All these layers stacked together gives a unified and complete view of a Docker image. These layers are compressed into a single image and can be pushed to the Docker Hub.

### Docker Hub

Docker Hub is a repository used by Docker to keep all Docker images in one central place. Docker images on Docker Hub can be private or public. Docker Hub also proposes automated builds and webhooks. Docker webhook is an HTTP call-back triggered by a *push* on external code repository. Automated builds, also known as autobuilds, automatically build Docker images from the source code located on external code repository (e.g. GitHub, BitBucket, etc.) and then host them automatically on Docker Hub. During the setup of Docker automated builds, a list of branches and tags, that the user wants to build into Docker images is created. When source code is pushed to a code repository (like GitHub) for these listed image tags, *push* uses Docker webhook to build a new Docker image [6]. The created image is then automatically pushed to Docker Hub. After the automated build, Docker sends HTTP request to a reachable Docker host on the Internet to notify it about the availability of new image, which pulls the built image and restarts container on the new image through Docker webhooks [7, 8].

### 1.1.2 Singularity

Singularity was introduced in 2016. It offers mobility of compute by facilitating portable environments through a single image file, which makes Singularity containers different from Docker's layered containers. Once Singularity image is created, it is hashed by using SHA256 hashing. Hence, it cannot be changed once it is created, consequently, it can be used for reproducing the results of scientific experiments. Singularity was introduced with the main goal of providing support for multi-user environment [9], where Docker failed to provide such support. Main features provided by Singularity are mobility of compute and reproducibility.

Mobility of compute is defined as creating and maintaining a workflow on a local machine that can be easily ported on other hosts, Linux distributions, and/or cloud service providers. To achieve this, Singularity packs everything that is required for the application to run in a distributable image. Then that image can be copied, archived, and shared. Moreover, Singularity containers are easily portable across different versions of the C library and different kernel implementations [10].

Same features of Singularity, which are used for mobility of compute, facilitate reproducibility as well. Hash of the image is also stored along with the image [10].

#### Singularity Images

Singularity takes snapshot, locks, and archives the developed application, which is known as Singularity image [11]. Due to image hashing, Singularity does not contain image layers, unlike Docker images. While publishing experimental results, authors can also publish Singularity image along with its hash, which allow other researchers to verify results.

**Singularity Hub**

Singularity Hub is a central repository provided by Singularity to keep Singularity images and their hash. It is designed to facilitate reproducible research and publications. Singularity Hub can be easily cited in the publications and hence other researchers can easily replicate conducted experiments. Singularity Hub connects with GitHub for automatic builds of Singularity images from the source code present at GitHub.

## 1.2 Thesis Outline

This thesis consists of four main parts divided into chapters. Chapter 2 provides literature survey on the security of containers. In accordance with that, it provides a list of security requirements that every OS-level virtualization technique should fulfill, and hence based on these requirements it investigates security of containers. This chapter also points out strong features and security weaknesses of containers. Chapter 3 presents various experiments done to analyze security of Docker and Singularity images used in two containerization frameworks used in neuroscience. Additionally, it provides results of these experiments. Finally, Chapter 4 concludes this thesis and highlights the future work.

# Chapter 2

# Security Review of Containers

In spite of the popularity of containers, there are various security concerns as well. Here, we focus our investigation on the security of Singularity and Docker containers because of three reasons. First, they are popular among containers. Second, security becomes a barrier in the adoption of containers in production. Third, both are already used in various environments, so it is easy to practically investigate their security. In this chapter, we study a set of security requirements, that each OS-level virtualization solution should provide. Additionally, this chapter investigates how these requirements are fulfilled by Singularity and Docker containers, and list their security implications. For this study, we categorize security into two broad categories: internal security provided by containers, and external security provided by the Linux operating system.

## 2.1 Container's Internal Security

We examine the security of containers based on the requirements given by Reshetova et al., 2014 [2] for comparing the security of a number of OS-level virtualization techniques. According to this research, an OS-level virtualization technique should satisfy the following requirements: process isolation, filesystem isolation, device isolation, IPC isolation, network isolation and limiting of resources inside containers. Apart

from that, in this section, we provide vulnerabilities that are added to containers by container images [12, 13].

### 2.1.1 Process Isolation

The main goal of process isolation is to prevent one container process from seeing or interfering with another container process. In other words, it limits the permissions and visibility of a container process to processes running in other containers. Containers achieve this goal through the use of namespaces in the Linux kernel. For this purpose, Docker uses *PID namespaces* which isolate a process with a particular process ID from the host and other containers. Provided that, it becomes difficult for an attacker to see processes running in other containers, hence harder to attack them [14]. In contrast to Docker, the default configuration of Singularity does not isolate PID namespaces of containers from the host. The reason behind that is Singularity's main goal is to provide mobility of compute and reproducibility, and not full isolation. However, this configuration can be changed to separate the *PID* namespaces either by using a command line or environment variable setting.

### 2.1.2 Filesystem Isolation

Filesystem isolation is required to prevent illegitimate access to filesystems of the host and containers. Docker uses *filesystem namespaces*, also called *mount namespaces*, for achieving this goal. Filesystem namespaces hide filesystems of the host and containers from other containers. However, some of the kernel filesystems are not namespaced, so Docker containers mount them for operation. For example, */sys, /proc/sys, /proc/sysrq - trigger, /proc/irq, and /proc/bus* are not namespaced [15]. This causes security concerns as Docker containers are directly able to access host filesystems. Consequently, Docker provides two filesystem protection mechanisms. First, Docker gives read-only permissions to containers for these filesystems. Second, containers are not allowed to remount any filesystem within containers. Additionally, Docker offers a *copy-on-write* filesystem mechanism, which allows containers to write to their specific filesystems and changes are not visible to other containers. Similarly,

files in Singularity containers are also isolated using *filesystem* namespaces.

### 2.1.3 Device Isolation

Applications and kernel access devices through special files known as device nodes. It is very crucial to limit the set of devices nodes that containers can access. This is primarily because an attacker can own the whole system by gaining access to some important device nodes such as /dev/mem, /dev/sd*, or /dev/tty. Docker uses the *Device Whitelist Controller* [16] feature of control groups to limit the set of devices that a Docker container can see and use. Additionally, Docker starts containers with *nodev* which prevents the use of already created device nodes inside the image. Furthermore, Docker does not allow containers to create new device nodes. However, some of the important device nodes cannot be namespaced such as */dev/mem, /dev/sd**, and kernel modules. Direct access to these device nodes by containers possess serious security concerns. Besides that, if a Docker container is executed in *privileged* mode then it gets access to all devices. Conversely, in default configuration of Singularity, all host devices are visible inside the container because the user is same inside and outside the container.

### 2.1.4 Inter-Process Communication (IPC) Isolation

IPC is a set of objects through which processes communicate with each other, such as shared memory segments, semaphores, and message queues. IPC isolation is needed to prevent containers from accessing or modifying data belonging to other containers, which is transmitted through these objects. Docker utilizes *IPC namespaces* to assign an IPC namespace to each container. Process in one IPC namespace cannot read or write IPC resources of another IPC namespace. The default configuration of Singularity does not provide IPC isolation.

### 2.1.5 Network Isolation

Isolation of container's network is very important to prevent network-based attacks such as address resolution protocol (ARP) spoofing and Man-in-the-Middle (MitM) Attack [17]. ARP spoofing is an attack which associates the attacker's Media Access Control (MAC) address with the Internet Protocol (IP) address of another host. ARP is a stateless protocol, hence hosts save all ARP replies even if they had not sent any ARP request for it [18], which becomes a source of attack. So, the host that has spoofed ARP response is not able to verify whether it belongs to legitimate host or attacker, and hence it starts sending packets at attacker's MAC address. In Singularity, by default, the container shares the network with the host. This is because Singularity tries to virtualize "as few as possible" namespaces. Whereas, Docker uses *network* namespaces [19] to isolate the network of Docker containers. Therefore, each Docker container has its different IP address, IP routing tables, network devices, etc. Consequently, Docker containers interact through the network interfaces with each other as well as with the host [20]. On the contrary, all containers share the same network bridge, which makes Docker vulnerable to ARP spoofing attack. In addition to that, bridge of Docker network forwards all incoming packets without any kind of filtering, which makes it vulnerable to Mac flooding [14]. In Docker, the network is used for two purposes : (i) to control Docker daemon remotely, (ii) to distribute images, which include pulling and saving images on Docker Hub.

#### Docker Daemon Remote Control

Docker daemon is controlled remotely through a socket (which is by default a Unix socket but can be changed to TCP socket), so Docker commands can be performed from another host remotely [21]. By accessing this socket, an attacker can pull and run containers in privileged mode, hence getting root access to the host.

#### Vulnerable Image Distribution Process

Docker images are present in compressed format on Docker Hub. So, Docker daemon pulls, uncompresses, and then runs containers from these images. Here, Docker daemon can be related to the package manager and Docker Hub to the software repository. Hence, Docker daemon also has an attack surface similar to the package manager. Vulnerabilities include storing, processing potentially untrusted code in Docker images by Docker daemon. The source code can be tampered during transfer or at the source. If some part of the network is compromised, an attacker can replace an image with malicious image, and that image gets downloaded on the host. As the image is in a compressed format, the attacker can cleverly craft the image (i.e. all zeros). If so, it has the potential of filling whole storage on the host after decompression, hence causing denial-of-service (DoS) attack. Other possible attacks are code injection or replay attacks. Additionally, the malicious image can be uploaded to the Docker Hub by an adversary. That image can be downloaded by millions of users infecting millions of machines.

For mitigation of these attacks, Docker introduced content trust which allows signing images before pushing to Docker Hub. However, this content trust can be disabled and thus disabling image signature check. Another issue is related to automated build and webhooks, where compromised GitHub account can lead to the execution of a malicious code. According to the experiment performed in [7] the malicious code was put in production in a very short time i.e within 5 minutes and 30 seconds of commit on GitHub. The content trust provides an environment where a single entity is trusted but in this case, trust is divided among several external entities.

### 2.1.6 Limiting of Resources

A DoS attack occurs when intended users are not able to use the system or network resources [22]. To launch DoS, the attacker floods targeted host or network traffic with superfluous requests to overload systems. Consequently, the target crashes or its resources get exhausted, hence disrupting normal execution of the system. To solve

this issue, both Docker and Singularity use *cgroups*. Cgroups restrict the amount of resources (CPU, memory, and disk I/O) that are used by containers, thus not allowing one container to consume all resources.

### 2.1.7 Vulnerabilities in Container Images

There can be vulnerabilities present inside the image itself when it is downloaded from the image repository. According to [23] over 30% official Docker images had high-priority common vulnerabilities and exposures (CVE) identifiers (IDs) and around 64% had high or medium level CVE vulnerabilities at the time of this work. This research work also states that Docker images with the *latest* tag also had vulnerabilities. These vulnerabilities are due to outdated packages contained in images, which may be a consequence of the use of old base image or due to pulling of outdated code during build.

Docker introduced Docker security scanning through which users can scan images to check whether they contain vulnerability or not. However, this scanning is limited to only private repositories and it is a paid service. The scan traverses all layers of the image and then identifies software packages in those layers. Further, it checks vulnerabilities in these software components by taking their Secure Hash Algorithms (SHAs) and then comparing against a standard list of CVEs. This scan can take up to 24 hours depending on image sizes. Additionally, this scanning technique does not detect malware, virus, or vulnerabilities which are not mentioned in the standard CVE database [24].

## 2.2 Security provided by Kernel

Out-of-the-box security provided by the Linux kernel, which secures the host from containers, is known as host hardening. Linux provides Linux capabilities and Linux Security Modules (LSM) to harden the security of the host system. Linux capabilities divide the privileges of superuser into pieces, and assign a subset of these capabilities

to specific processes. Whereas, LSM provides a framework for Linux to support various security modules. Currently, three security modules are officially integrated with Linux kernel which includes SELinux, AppArmor, and Seccomp. Out of these three, only the first two are supported by Docker. Docker only integrates with Seccomp if LXC are used. Below, we provide the details of these host hardening techniques.

### 2.2.1 Linux Capabilities

According to the official documentation of Linux [25], Unix systems traditionally categorized processes as *privileged* processes and *unprivileged* processes. Privileged processes are root users with zero user id (UID), whereas unprivileged users are normal users with nonzero UID. Privileged processes are exempted from permission check, whereas unprivileged processes are liable to full permission checks. Linux divides privileges of superuser into different pieces, known as capabilities, which can be independently enabled or disabled. As a result, Docker can disable some of the capabilities of containers, thus improving the security of the system. As Docker containers share the kernel with the host, most of their tasks are done by the host. As a consequence, disabling some of the capabilities in Docker containers do not affect their functionality. For example, *CAP_NET_ADMIN* capability allows configuration of the network, which can be disabled in Docker containers because all network configurations are handled by Docker daemon. By default, most of Linux capabilities are disabled when Docker container is started in order to secure the host system from attackers [1]. Table 1 lists some of the capabilities that are disabled in Docker containers.

| | |
|---|---|
| CAP_SETPCAP | Modify process capabilities |
| CAP_SYS_MODULE | Insert/Remove kernel modules |
| CAP_SYS_RAWIO | Modify kernel memory |
| CAP_SYS_PACCT | Configure process accounting |
| CAP_SYS_NICE | Modify priority of processes |
| CAP_SYS_RESOURCE | Override resource limits |
| CAP_SYS_TIME | Modify the system clock |
| CAP_SYS_TTY_CONF IG | Configure tty devices |
| CAP_AUDIT_WRITE | Write the audit log |
| CAP_AUDIT_CONTROL | Configure audit subsystem |
| CAP_MAC_OVERRIDE | Ignore kernel MAC policy |
| CAP_MAC_ADMIN | Configure MAC configuration |
| CAP_SYSLOG | Modify kernel printk behavior |
| CAP_NET_ADMIN | Configure the network |
| CAP_SYS_ADMIN | Catch all |

Table 1: Some disabled capabilities in Docker containers [1].

### 2.2.2 Linux Security Modules

Below, we provide details of two LSMs that are currently supported by Docker.

#### SELinux

SELinux is security enhancement to the Linux system, which integrates Mandatory Access Control (MAC). MAC strongly separates all applications, which in turn decreases potential damage if an application is compromised. SELinux classifies active users or processes as subjects, and all system resources as objects. Everything in SELinux is controlled by labels [26]. System administrator writes SELinux policies, which control accesses of system objects by processes. These policies are of three types: Type Enforcement (TE), Multi-Category Security (MCS), and Multi-Level

Security (MLS). Out of these three, Docker uses only the first two types of policies [14]. Using TE, Docker protects the host from containers, by forcing container processes to read/write content that has a specific label [14]. MCS is used to protect containers from other containers. To achieve that, each container is assigned a unique MCS label. All the files that belong to a container are also labelled with the same MCS label. Therefore, kernel does not allow a container process to read/write to a file that have different MCS label. Hence isolating all containers.

### AppArmor

Like SELinux, AppArmor is another MAC solution for enhancing security of Linux systems. It uses the concept of file system paths instead of labels. AppArmor mentions the file path in the binary of the application along with the allowed permissions on that file. Two modes are supported by AppArmor: enforcement mode and complain/learning mode [14]. Enforcement mode is used to enforce policies that are defined in the AppArmor profile, whereas complain/learning mode also allows violations of the profile policies. However, these violations are logged, and may be used later for developing new profiles. Docker provides an interface to systems that support AppArmor. This interface is used to load AppArmor profile for containers. Consequently, when containers are launched, a pre-defined AppArmor profile is loaded automatically, if administrator does not specify an AppArmor profile. These security profiles are loaded in enforcement mode by default, to make sure that the policies defined in the profile are enforced. Therefore, the important files of the host such as */sys/fs/cgroups/* and */sys/kernel/security/* remain protected from containers.

### Limitations of LSMs

Security provided by these modules is limited due to the generic nature of profiles provided by these security modules [7]. For example, default SELinux profile assigns same domain to all Docker containers, which helps in protecting host from containers, but not containers from containers. Similarly, default AppArmor profile provide full access of network system and capabilities to Docker containers. To improve the

security of the host, a potential solution is to write container-specific profiles.

## 2.3 Conclusion

In this chapter, we examine various security threats that containers are prone to. We focus our investigation on the security of Docker and Singularity containers because of their popularity. In that context, we study a list of security requirements that every OS-level virtualization should fulfill. Later, we investigate security of Docker and Singularity containers based on these requirements. We broadly categories security of containers into two types: internal security provided by containers, and external security provided by the Linux kernel. Internal security includes process isolation, device isolation, IPC isolation, network isolation, limiting of resources in containers, and vulnerabilities in images. We then highlight solutions provided by containers to provide internal security, and later, we discuss how efficient these solutions are. Further, we discuss external security, which is security provided by the Linux kernel, to protect the host from containers, or containers from containers. External security includes Linux capabilities and LSM framework. Linux capabilities can be disabled/enabled as needed, and LSM framework is used to define various security profiles for containers. Inspite of all these efforts, it is clear that there is still a need of security enhancement in containers.

# Chapter 3

# Image Analysis

In this chapter, we analyze container images, which are used in two neuroscience frameworks. In accordance with that, these images are scanned using four popular scanners. Additionally, two different approaches, *image update* and *image minification* are followed in order to see their effect on the security vulnerabilities. Finally, this chapter discuss results of these experiments.

## 3.1 Image Scanning

Image scanning is a technique to detect image vulnerabilities, which reduces the risk of vulnerability exploitation inside containers. Scanning refers to the practice of collecting all information about the container image, investigating it for vulnerabilities, and finally producing a report to summarize scanning results. These reports can be used for security assessments.

### 3.1.1 Container Images

We scanned all container images available at the time of this study on two containerization frameworks used in neuroscience: BIDS apps [27] (26 images) and Boutiques [28] (18 images), totalling 44 container images (Table 2). At the time of the

study, BIDS apps had 27 images, out of which one wasn't available on DockerHub. Boutiques had 49 images, however, only 23 unique images were listed, out of which 3 couldn't be retrieved and 2 were already included in BIDS apps. All the final 26 images from BIDS apps were Docker images, whereas the 18 Boutiques images contained 12 Docker images and 6 Singularity images.

### 3.1.2 Image Scanners

We used four container image scanners : Anchore, Vuls, and Clair to scan Docker images, and Singularity Container Tools (Stools) to scan Singularity images.

#### Anchore

Anchore is an end-to-end, open-source container security platform. It analyzes container images and lists vulnerable OS packages, non-OS packages (Python, Java, Gem, and npm), and files. In our experiments, we used Anchore Engine version 0.5.0 through Docker image `anchore/anchore-engine:v0.5.0`, and Anchore vulnerability database version 0.0.11.

| **Abbrv** | **Image** | **Distribution** |
|---|---|---|
| k | bids/hyperalignment | ubuntu:16.04 |
| l | bids/niak | ubuntu:16.04 |
| h | bids/fibredensityandcrosssection | ubuntu:14.04 |
| g | bids/ndmg | ubuntu:14.04 |
| f | bids/ndmg:v0.1.0 | ubuntu:14.04 |
| j | bids/oppni:v0.7.0-1 | ubuntu:14.04 |
| Q | bids/brainiak-srm | ubuntu:16.04 |
| e | bids/tracula:v6.0.0-4 | ubuntu:14.04 |
| H | bids/rs_signal_extract:0.1 | ubuntu:16.04 |
| d | bids/example | ubuntu:14.04 |
| i | bids/cpac:v1.0.1a_22 | ubuntu:16.04 |
| S | bids/mindboggle:0.0.4-1 | debian:8 |

| | | |
|---|---|---|
| F | bids/nipypelines:0.3.0 | debian:8 |
| N | bt5e/ants:latest | centos:7 |
| a | poldracklab/fmriprep:1.2.3 | ubuntu:16.04 |
| Z | bids/hcppipelines:v3.17.0-18 | debian:8 |
| b | poldracklab/fmriprep:unstable | ubuntu:16.04 |
| c | bids/dparsf:v4.3.12 | ubuntu:14.04 |
| Y | poldracklab/mriqc:0.15.0 | ubuntu:16.04 |
| x* | shots47s/bids-fmriprep-1.2.3 | ubuntu:16.04 |
| R | gkiar/dwipreproc_fsl-5.0.11_minified | ubuntu:16.04 |
| U | bids/freesurfer | ubuntu:14.04 |
| I | bids/antscorticalthickness:v2.2.0-1 | ubuntu:17.04 |
| V | bids/mrtrix3_connectome | ubuntu:18.04 |
| O | bigdatalabteam/hcp-prefreesurfer:exec-centos7-fslbuild-centos5-latest | centos:7 |
| P | bigdatalabteam/hcp-prefreesurfer:exec-centos7.freesurferbuild-centos4-latest | centos:7 |
| v* | aces/cbrain-containers-recipes:fsl_v6.0.1 | ubuntu:16.04 |
| D | bids/broccoli:v1.0.0 | centos:6 |
| w* | shots47s/bids-freesurfer-6.0 | ubuntu:14.04 |
| s* | c3genomics/genpipes | centos:7 |
| W | bids/spm | ubuntu:14.04 |
| X | bids/aa:v0.2.0 | ubuntu:14.04 |
| E | mcin/docker-fsl:latest | centos:7 |
| K | mcin/ica-aroma:latest | centos:7 |
| G | bids/baracus | ubuntu:14.04 |
| T | camarasu/creaphase:0.3 | centos:7 |
| M | mcin/qeeg:latest | centos:7 |
| L | bids/magetbrain | ubuntu:18.04 |
| u* | MontrealSergiy/BEst | ubuntu:16.04 |
| J | bids/afni_proc | ubuntu:17.1 |
| C | gkiar/mask2boundary:v0.1.0 | alpine:3.9.0 |
| t* | bioinformatics-group/aqua-singularity-recipe | debian:9 |

| | | |
|---|---|---|
| A | gkiar/onevoxel:v0.3.0rc2 | alpine:3.7.1 |
| B | bids/rshrf:1.0.1 | alpine:3.8.4 |

Table 2: List of images scanned. Image with asterisk (*) mark are Singularity image and others are all Docker images.

### Vuls

Vuls is an open-source vulnerability scanner for Linux and FreeBSD. It offers both static and dynamic scanning, and both local and remote scanning. In our experiments, we used Vuls 0.9.0, executed through Docker image `vuls/vuls:0.9.0` in remote dynamic mode.

### Clair

Clair is an open-source and extensible vulnerability scanner for Docker and appc container images, developed by CoreOS (now Container Linux), a Linux distribution to deploy container clusters. Clair has a client-server architecture, in which the server scans Docker images layer by layer and maintains a database of vulnerabilities. We used Clair through Clair-scanner, a tool to facilitate the testing of container images against a local Clair server. lair-scanner scans the image, prepares a list of vulnerabilities, compares that list against a whitelist, and flags vulnerabilities that are not present in the whitelist. In our experiments, we did not use a whitelist to filter scanning results in order to make a fair comparison between scanners. Clair-scanner maintains a Docker image with the up-to-date vulnerability database from a set of different sources. We used Clair version 2.0.6, executed through Docker image `arminc/clair-local-scan:v2.0.6`. For the vulnerability database, we used Docker image `arminc/clair-db:latest`, last updated on 2019-09-18.

#### Stools

Singularity Tools (Stools) are an extension of Clair for Singularity images. Stools exports Singularity images to `tar.gz` format, acting as a single layer Docker image to circumvent the Docker-specific requirements in the Clair API. In our experiments, we used Singularity Tools version 3.2.1 through Docker image `vanessa/stools-clair:v3.2.1`. Since Stools uses Clair internally for scanning, vulnerability databases used by Stools are same as mentioned for Clair. To scan Singularity images, we followed the steps mentioned in the Stools documentation.

### 3.1.3 Vulnerability Databases used by Scanners

Scanners refer to two types of vulnerability databases (Table 3). The first one is the Open Vulnerability and Assessment Language (OVAL) database, an international open standard that supports various OS distributions including Ubuntu, Debian and CentOS but not Alpine. The second one are vulnerability databases from specific OS distributions, such as Alpine-SecDB, Debian Security Bug Tracker, Ubuntu CVE Tracker, or Red Hat Security Data. In these databases, OS distributions often assign a status to each vulnerability, to keep track of required and available security fixes in different versions of the distribution. Vuls uses OVAL databases for all distributions except Alpine, whereas Anchore uses OVAL only for CentOS. On the contrary, Clair exclusively refers to distribution-specific databases, as distribution-specific databases are assumed to be more complete. It is also worth noting that there is no vulnerability data for Ubuntu 17.04 and 17.10 distributions in the OVAL database, since these distributions have reached end of life, meaning that images with these distributions cannot be scanned with Vuls. For CentOS images, Anchore and Clair give scanning results using Red Hat Security Advisory (RHSA) identifiers, whereas Vuls uses the Common Vulnerabilities and Exposures (CVE) identifiers used in OVAL. We mapped RHSA identifiers to corresponding CVE identifiers, to allow for a comparison between scanners. Also, all these scanners refer to National Vulnerability database (NVD), which is a vulnerability database launched by the National Institute of Standards and Technology (NIST), in order to get additional information about the vulnerabilities (e.g., its severity scores, its fix information, and its impact ratings).

Different vulnerabilities may be reported by scanners if scanning experiments take place on different dates. To avoid such discrepancies, we froze the vulnerability databases used by these scanners as of 2019-09-25.

| OS | Anchore | Vuls | Clair |
|---|---|---|---|
| **Alpine** | Alpine-SecDB | Alpine-SecDB | Alpine-SecDB |
| **CentOS** | Red Hat OVAL Database | Red Hat OVAL Database and Red Hat Security Advisories | Red Hat Security Data |
| **Debian** | Debian Security Bug Tracker | Debian OVAL Database and Debian Security Bug Tracker | Debian Security Bug Tracker |
| **Ubuntu** | Ubuntu CVE Tracker | Ubuntu OVAL Database | Ubuntu CVE Tracker |

Table 3: Vulnerability databases used by scanners for different OS distributions. All scanners also refer to the National Vulnerability Database (NVD) for vulnerability metadata.

## 3.2 Detected Vulnerabilities

Table 4 shows results of scanning BIDS and Boutiques images with four scanners. An important amount of vulnerabilities were found in the tested container images, with an average of 460 vulnerabilities per image and a median of 321. Moreover, a significant fraction of detected vulnerabilities are of high severity (CVSS score <= 7.0) and a few of them are of critical severity (CVSS >= 9.0) (Fig 1). Remote attackers could possibly exploit these vulnerabilities to execute arbitrary code in the container, by crafting responses to specific network requests. Images based on the Alpine distribution had the lowest numbers of vulnerabilities, but no significant difference in the numbers of vulnerabilities detected in Ubuntu, Debian or CentOS

distributions was observed.

Unsurprisingly, a strong linear relationship is found between the number of detected vulnerabilities and the number of packages present in the image (Fig 4). On average, 1.7 vulnerabilities are introduced for each new package installation. This observation motivates a systematic review of software dependencies by application developers, to avoid unnecessary packages in container images. This is also an argument in favor of lightweight distributions such as Alpine. Compared to Ubuntu and Debian distributions, CentOS images seem to have a lower number of vulnerabilities by package on average, although data is too scarce to conclude.

All the collected data are available in our GitHub repository at https://github.com/big-data-lab-team/container-vulnerabilities-paper with a Jupyter notebook to regenerate the figures.

| **Abbrv** | **Anchore** | **Clair** | **Vuls** | **Stools** |
|---|---|---|---|---|
| k | 1717 | 1327 | 1433 | None |
| l | 1482 | 1366 | 1420 | None |
| h | 1086 | 849 | 868 | None |
| f | 934 | 758 | 782 | None |
| g | 934 | 757 | 782 | None |
| j | 913 | 698 | 723 | None |
| Q | 897 | 484 | 503 | None |
| e | 874 | 664 | 683 | None |
| H | 867 | 454 | 473 | None |
| d | 838 | 663 | 683 | None |
| i | 698 | 524 | 535 | None |
| S | 671 | 955 | 569 | None |
| F | 641 | 878 | 451 | None |
| N | 592 | 600 | 336 | None |
| a | 591 | 441 | 447 | None |
| Z | 526 | 873 | 620 | None |
| b | 516 | 355 | 393 | None |
| c | 507 | 750 | 757 | None |

| | | | | |
|---|---|---|---|---|
| Y | 432 | 328 | 332 | None |
| R | 391 | 258 | 263 | None |
| U | 353 | 215 | 192 | None |
| I | 289 | 27 | None | None |
| V | 275 | 157 | 158 | None |
| O | 257 | 266 | 157 | None |
| P | 257 | 266 | 157 | None |
| D | 203 | 203 | 117 | None |
| W | 191 | 246 | 235 | None |
| X | 189 | 244 | 233 | None |
| E | 178 | 178 | 201 | None |
| K | 178 | 178 | 202 | None |
| G | 172 | 238 | 216 | None |
| T | 148 | 148 | 186 | None |
| M | 121 | 121 | 142 | None |
| L | 112 | 117 | 126 | None |
| J | 67 | 16 | None | None |
| C | 15 | 11 | 11 | None |
| A | 5 | 2 | 2 | None |
| B | 5 | 3 | 3 | None |
| x* | None | None | None | 421 |
| v* | None | None | None | 206 |
| w* | None | None | None | 203 |
| s* | None | None | None | 197 |
| u* | None | None | None | 105 |
| t* | None | None | None | 7 |

Table 4: Number of vulnerabilities detected by four different scanners. Singularity Images are only scanned with Stools.

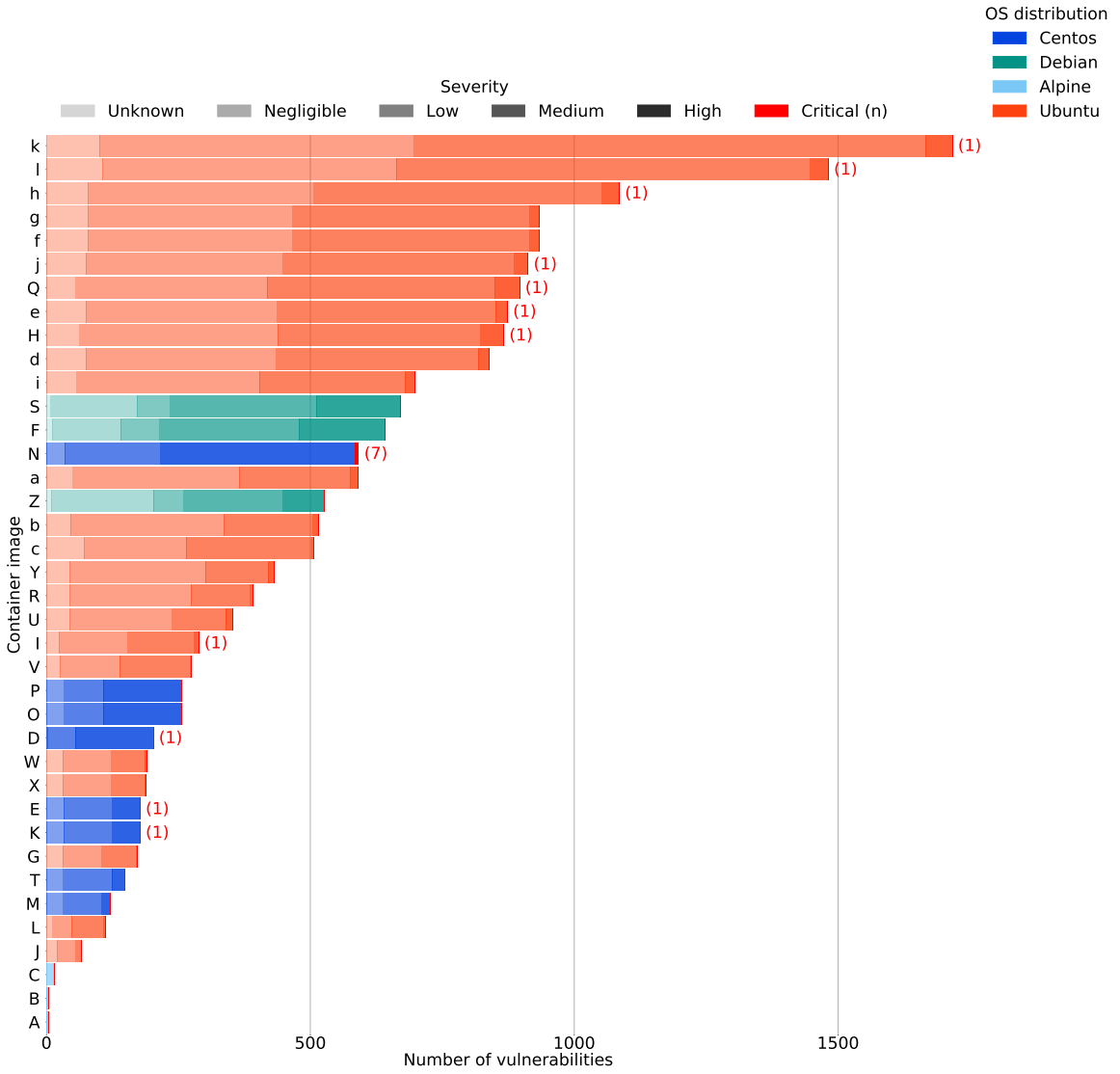


Figure 1: Number of vulnerabilities detected in tested container images. Number of vulnerabilities by container image and severity, showing hundreds of detected vulnerabilities per image. Images s*,t*,u*,v*,w* and x* are Singularity images scanned by Stools and others are Docker images scanned using Anchore.

## 3.3 Differences between Scanners

Important discrepancies were found between scanners (Fig 2), in particular between Anchore and the other two scanners, for which Jaccard coefficients as low as 0.6 were

found, meaning that scanning results only overlapped by 60%. Vuls and Clair appear to be in better agreement, with a Jaccard coefficient of 0.8. The question arises why these scanners report different number of vulnerabilities. How their logic differs and what exactly is making these results different than each other.

We discuss these discrepancy reasons later in this section. Before that, we discuss Ubuntu Vulnerability Databases that are needed to understand the discrepancy reasons.

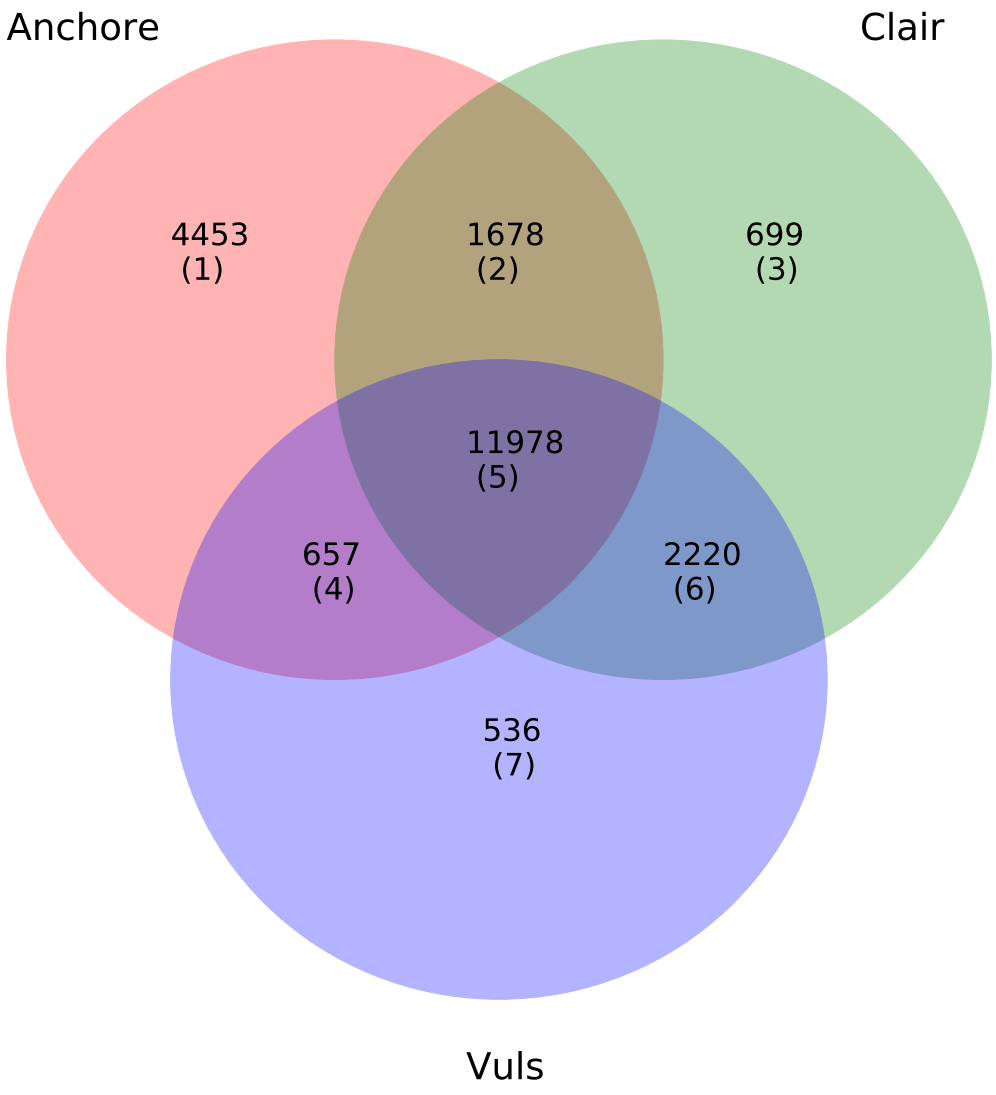

Figure 2: Differences between vulnerabilities detected by the different scanners. The Jaccard coefficients between the sets of detected vulnerabilities are quite low, showing important discrepancies between the scanners: Jaccard(Anchore, Clair) = 0.63, Jaccard(Anchore, Vuls) = 0.59, Jaccard(Vuls, Clair) = 0.80. Two Ubuntu 17.04 images weren't included in this comparison as they cannot be scanned by Vuls.

### 3.3.1 Ubuntu Vulnerability Databases

Through our experiments, we find that image scanners report vulnerabilities differently depending on the different Linux release types. So, it is important to provide an overview of these Linux releases types. BIDS app and Boutiques images in our experiments have diverse Linux distributions, such as Ubuntu, Centos, Debian, and Alpine. Further, a particular distribution has different releases in general. To illustrate, we here examine Ubuntu as an example, since it is heavily utilised in our experiments. However, other distributions also have similar release trend. Ubuntu releases can be of two types: Regular or Long Term Support (LTS). Regular release is supported only for 9 months, whereas LTS is supported for 5 years. Once LTS support reaches End-Of-Life (EOL), Ubuntu decides whether to provide extended security support or not, which comes in the form of Extended Security Maintenance (ESM) release. ESM is a paid service that can be requested from Ubuntu to get security updates even after the EOL of a particular Ubuntu release. ESM and LTS releases of a particular Ubuntu distribution have different vulnerability databases, because after EOL of a release, a vulnerability is fixed only in packages of ESM; however, it still exists in LTS.

As an example, Ubuntu has ended support for some releases, such as 14.04 LTS, 17.04, and 17.10, as they reached EOL. However, Ubuntu decided to provide optional extended support for Ubuntu 14.04. To use this extended support, users have to purchase ESM release for Ubuntu 14.04. The images in our experiments have Ubuntu 14.04 LTS, so scanners should refer to LTS release only while collecting data from the vulnerability databases. More interestingly, scanners are able to refer to ESM release vulnerability database even without purchase, however, this may lead to discrepancies in results. Further, if an image is using a Linux release where EOL has

been reached that image may turns to be vulnerable as it is not possible to retrieve any updates.

Ubuntu maintains its vulnerability database in the Ubuntu Launchpad or the Ubuntu CVE Tracker. For a given CVE, there is an entry for all ubuntu releases. Ubuntu releases are listed with the package name in which vulnerability exists and are entitled a status (https://git.launchpad.net/ubuntu-cve-tracker/plain/README), which is encoded in the following form:

`<release>_<source-package>: <status> (<version/notes>)`

For a given release, the status can be any one of the following:

**DNE (Does Not Exist):** The package does not exist in the archive.

**needs-triage:** The vulnerability of this package is not known, and hence evaluation is needed.

**not-affected:** The package, while related to the CVE in some way, is not affected by the issue. The <notes> should provide detailed information, if needed. For instance, if the given status is `not-affected (1.14.6-1.1)` then this indicates that this specific package version is not affected.

**needed:** The package is vulnerable to the CVE and needs to be fixed.

**active:** The package is vulnerable to the CVE, needs fixing, and is actively being worked on.

**ignored:** The package, while related to the CVE in some way, is being ignored for some reason. The <notes> should provide that reason. For instance, if status looks like `ignored (reached end-of-life)`, this means that the given Ubuntu release already reached end-of-life, so this CVE is ignored by Ubuntu. However, if there is extended security support for this release then this CVE will be handled in ESM release.

**pending:** The package is vulnerable and has been fixed but an update has not been yet uploaded or published. The <version> is given to indicate the particular version

where the fix has been done.

**deferred:** The package is vulnerable, but its fix has been deferred for some reason. The <notes> should provide further details. If a date is mentioned, e.g. `deferred (2015-02-02)`, the given date specifies the date when the CVE was deferred.

**released:** The package was vulnerable, but an update has been already uploaded and published, e.g. `released (1.2.3)`, <version> indicates the first version where the fix was applied.

**released-esm:** The package was vulnerable and an update has been already uploaded and published. However, this update is published in the Ubuntu ESM release only and not in LTS. The fixed version of such packages is appended by either +esm or ~esm, indicating that this package version is available only via ESM.

Figure 3 illustrates how the information of a CVE is represented in the Ubuntu launchpad database.

```
Candidate: CVE-2012-5340
PublicDate: 2020-01-23 22:15:00 UTC
References:
 https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2012-5340
 https://git.ghostscript.com/?p=mupdf.git;a=commitdiff;h=f919270b6a732ff45c3ba2d0c105e2b39e9c9bc9 (1.1)
 http://www.exploit-db.com/exploits/23246
Description:
 SumatraPDF 2.1.1/MuPDF 1.0 allows remote attackers to cause an Integer
 Overflow in the lex_number() function via a corrupt PDF file.
Ubuntu-Description:
Notes:
Mitigation:
Bugs:
Priority: medium
Discovered-by:
Assigned-to:
CVSS:
 nvd: CVSS:3.1/AV:L/AC:L/PR:N/UI:R/S:U/C:H/I:H/A:H


Patches_mupdf:
upstream_mupdf: released (1.2-2)
precise/esm_mupdf: DNE
trusty_mupdf: ignored (out of standard support)
trusty/esm_mupdf: DNE
xenial_mupdf: needs-triage
bionic_mupdf: not-affected (1.12.0+ds1-1)
eoan_mupdf: not-affected
focal_mupdf: not-affected
groovy_mupdf: not-affected
devel_mupdf: not-affected
```

Figure 3: Example of a CVE representation in Ubuntu Launchpad Database

### 3.3.2 Discrepancy Reasons

According to our investigation of these results precisely, below are some of the reasons of getting different results from these scanners. These reasons include bugs in scanner's logic, some scanners ignoring vulnerabilities in a particular package, referring to different databases, how frequently they are updating their database, etc,. We have explained these differences below in detail.

1. **Sub-packages vulnerabiities-** Some vulnerabilities in region **1** are found in sub-packages of vulnerable packages: They are correctly reported by Anchore and missed by Vuls and Clair.

2. **Linux-libc-dev Package-** Anchore is detecting vulnerabilities in the development package of the C library (`linux-libc-dev` in Ubuntu and Debian). Clair is only detecting Debian vulnerabilities in `linux-libc-dev`, whereas Vuls is not detecting vulnerabilities in `linux-libc-dev` at all. This is the reason behind vulnerabilities in region **1** and region **2** of the Figure 2.

3. **Vulnerabilities in ESM-** These are the vulnerabilities which are present in Ubuntu 14.04 ESM release. These vulnerabilities are incorrectly missed by Anchore and Vuls: they have been detected in ESM but were already present in LTS. This reason is the main reason for vulnerabilities in the region **3** of the Figure 2.

4. **Database difference-** Clair is using a Docker image for database. So it is not updated everyday. When we scanned images we used the latest available container. For example, CVE-2019-5094 was not updated in Clair container that time. This reason adds to region **3** of the Figure 2.

5. **Vulnerabilities not in ESM-** Vulnerabilities, which are not present in Ubuntu 14.04 ESM, are not detected by Clair but they affect Ubuntu 14.04 LTS, which is present in the image. This is the main reason behind region **4** of the Figure 2.

6. **Epoch bug-** There is a bug in the Anchore logic due to which some vulnerabilities are not detected by Anchore. This bug gets triggered when package's version have epoch (1:3.3.9-1ubuntu2.3) in it. This is one of the major reason

behind region **6** of the Figure 2.

7. **Debian's minor vulnerabilities-** Vulnerabilities marked `minor` by Debian are ignored by Anchore. However, it is reported by Clair and Vuls. This is also one of the major reason behind region **6** of the Figure 2 in addition to *epoch bug*.

8. **Out of standard support bug-** There is another bug in Anchore due to which it is not able to detect some vulnerabilities. Due to this bug Anchore does not detect vulnerabilities which have a status as `ignored (out of standard support)` in Ubuntu 14.04 LTS. This is also one reason behind region **6** of the Figure 2.

9. **Ignored (reached end-of-life)-** Some vulnerabilities have status as `ignored (reached end-of-life)` so Anchore checks the last status of these vulnerabilities whether OS package was vulnerable to the cve prior to the end-of-life. If it was vulnerable only then Anchore will report it, otherwise not. However, Clair may or may not detect it depending on its vulnerability status in ESM release. If the vulnerability status is `DNE( Does Not Exist)` in ESM, Clair does not detect it. Accordingly, this reason adds vulnerabilities to the region **6** and **3** of the Figure 2.

10. **Rejected CVEs-** There are some CVEs that are rejected because they are duplicate copies of another CVEs. Only Vuls is reporting these rejected vulnerabilities. This is one of the reasons behind vulnerabilities in the region **7** of the Figure 2.

11. **Debian's Temporary vulnerabilities-** Vulnerabilities that are flagged temporary by the Debian distribution are reported by Vuls but not by Anchore or Clair. This is also one of the reasons that adds to the region **7** of the Figure 2.

12. **Misses by Clair and Anchore-** Clair and Anchore are missing vulnerabilities in Centos images. We weren't able to explain why they were detected by Vuls only. This is the major reason behind region **7** of the Figure 2.

## 3.4 Approaches to Reduce Vulnerabilities

In comparison to tested images, no vulnerabilities were found in latest updated base Docker images `ubuntu:20.04` and `centos:7`. So we decided to update these images and see the effect on the number of vulnerabilities. Updating an image involves updating all software packages that are inside that image. However, this option effects the reproducibility of the image, which means that the image cannot be used to verify research findings of other scientists. An alternative approach can be to trim unnecessary packages from images, which in turn reduces the number of vulnerabilities present in images. Therefore, in this section, we discuss these two approaches and their effect on the vulnerabilities.

### 3.4.1 Image Update

A first approach to reduce the number of vulnerabilities in container images is to update their packages to the latest version available in the OS distribution. To study the effect of such updates, we developed a script (available here) to identify the package manager in the image, and invoke it to update all OS packages. We updated images on 2019-11-05.

### 3.4.2 Image Minification

A second approach to reduce the number of vulnerabilities in the images is to remove unnecessary packages, an operation potentially specific to each analysis. We used the open-source ReproZip tool [29] to capture the list of packages used by an analysis. ReproZip first captures the list of files involved in the analysis, through system call interception, then retrieves the list of associated software packages, by querying the package manager. We extend this list with a passlist of packages required for the system to function, such as `coreutils` and `bash`, and with all the dependencies of the required packages, retrieved using Debtree. Repoquery could be used in RPM-based distributions instead. Our minification script, available here, installs ReproZip in the image to minify, runs an analysis to collect a ReproZip trace, and finally deletes

all unnecessary packages. We had used the Neurodocker tool initially, but it did not affect the detected vulnerabilities as it was removing unused files without using the package manager.

Using this approach, we minified five Debian- or Ubuntu-based BIDS app images, using basic analysis examples found in the applications documentation.

### 3.4.3 Effect of image update

Updating container images reduces the number of vulnerabilities by package by a factor of 3 on average, resulting in only 0.6 extra vulnerabilities by package (Fig 4, r=0.81, $p<10^{-8}$). Twelve container images are missing on this figure: six of them could not be updated due to various issues with the package manager, and six of them are Singularity images that we didn't update. In spite of the associated reproducibility challenges, updating packages therefore appears to be an efficient way to avoid vulnerabilities. It is not an ultimate solution though, as a substantial number of vulnerabilities remain.

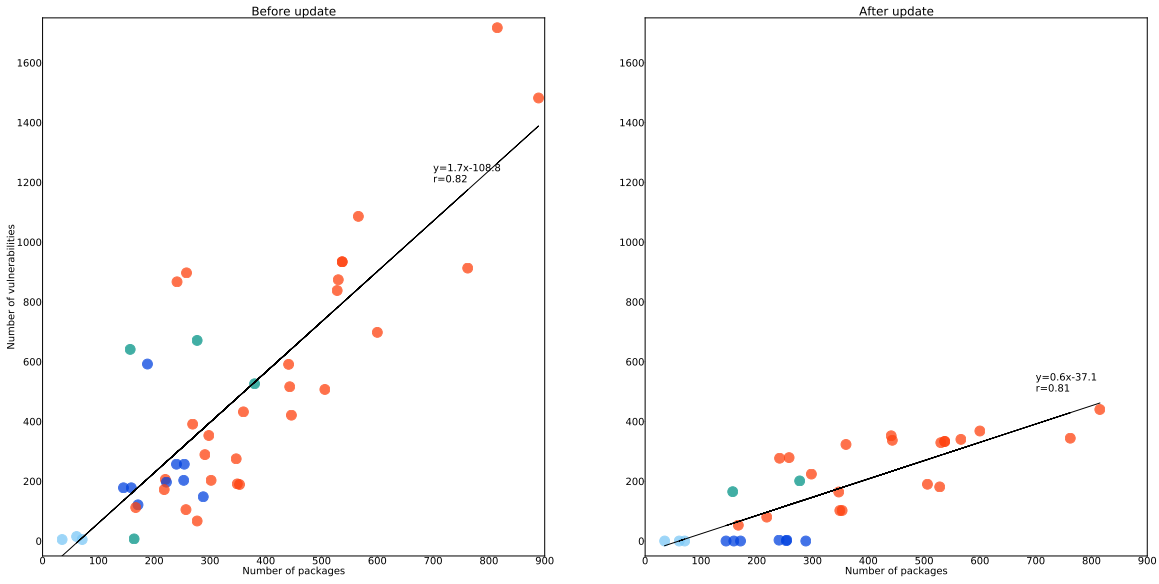


Figure 4: Number of vulnerabilities by number of packages, showing a strong linear relationship.

### 3.4.4 Effect of minification

Another approach to reduce the number of vulnerabilities involves deleting unnecessary packages from the container images. It is a tedious operation, as it requires running an actual data analysis in the container image, to identify the packages required by the application. In addition, the resulting container image is only valid for the specific type of analysis used in the minification process, as other executions might require a different set of packages.

Using the ReproZip-based approach described previously, we minified 5 different images covering the spectrum of detected vulnerabilities (Fig 5). We find that minification reduces the number of vulnerabilities, albeit less systematically than package update. For some container images, such as image **S**, minification removes more than 70% of the detected vulnerabilities. For other images, such as image **g**, it only reduces the number of vulnerabilities by less than 1%. The effect of minification stems from the number of packages that can be removed, which varies greatly across images. For instance, images **g** and **a** have a large number of packages, but the last majority of them is required by the analysis, which makes minification less useful. In other cases, a limited number of unnecessary packages contain a significant number of vulnerabilities, which makes minification very impactful. This was the case in images **d**, **S** and **U**, where removing compilers and kernel headers reduced the number of vulnerabilities by an important fraction.

### 3.4.5 Combined effect of image update and minification

Package update and image minification remove different types of vulnerabilities. The former is efficient against vulnerabilities that have been fixed by package maintainers, while the latter targets unused software. In two of the five tested images (images **S** and **U**), we find that combining update and minification further reduces the number of vulnerabilities compared to using only one of these processes (Fig 5).

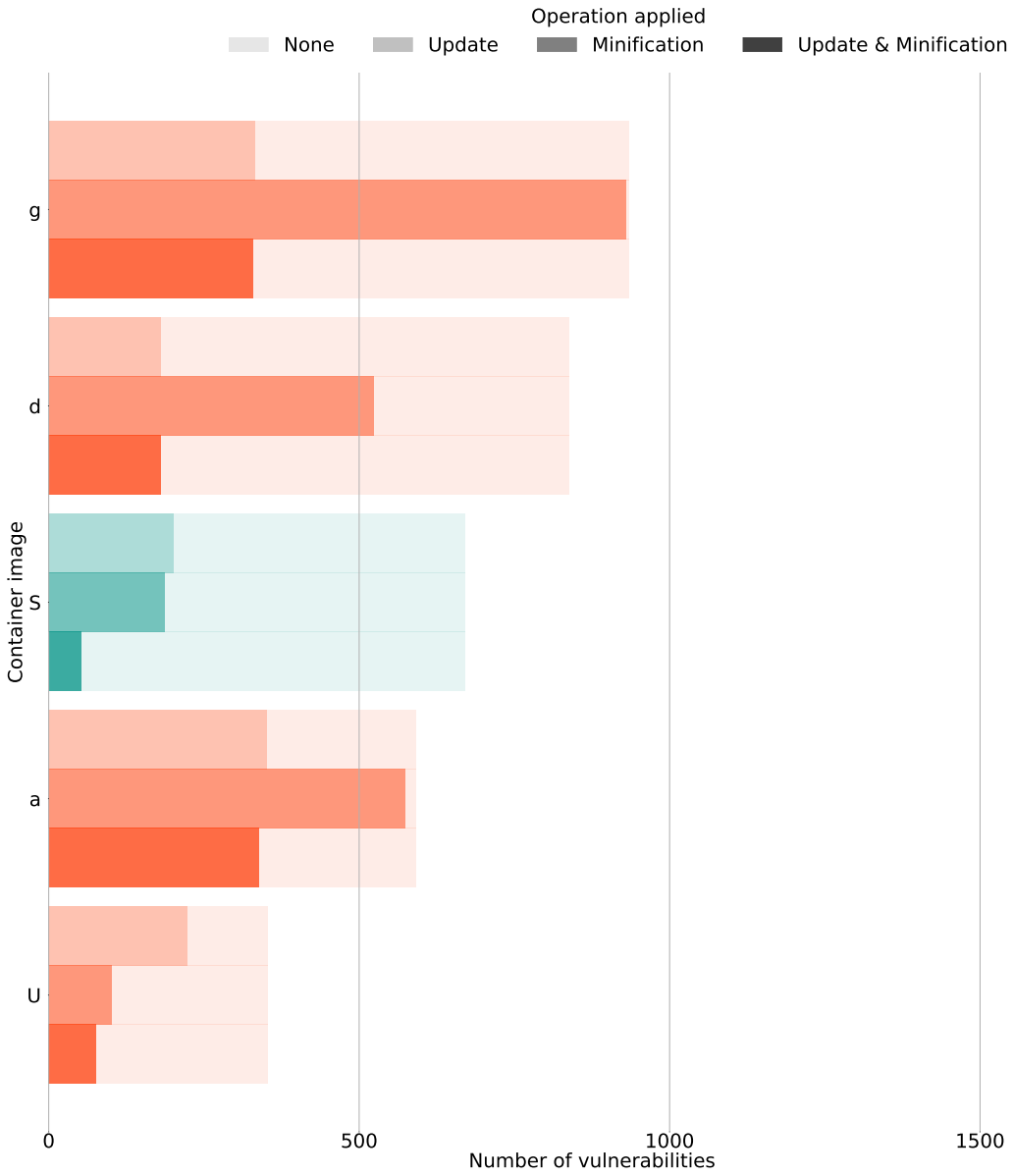


Figure 5: Effect of image minification and package update on 5 container images, showing that both techniques are complementary.

# Chapter 4

# Conclusions

This chapter discusses the general conclusions, contributions, and the future work.

## 4.1 General Conclusions

We conclude that there is a widespread issue with security vulnerabilities in container images used for neuroimaging analyses, and it is likely to impact other scientific disciplines as well. As shown in our results, it is common for container images to hold hundreds of vulnerabilities, including several of critical severity. Container images are impacted regardless of the type of analyses that they support, and the main OS distributions Ubuntu, Debian and CentOS are all affected. This issue can be addressed by the application software developers by: (1) minifying container images, by using lightweight OS distributions and reducing software dependencies, and (2) applying regular security updates, which requires using OS distributions with long-term support.

Software updates remove about two-thirds of the vulnerabilities found and should certainly be considered the primary solution to this problem. However, in neuroimaging as in other disciplines, software updates are generally discouraged because they can affect analysis results by introducing numerical perturbations in the computations [30, 31]. We believe that this position is not viable from an IT security

perspective, and that it could endanger the entire Big Data processing infrastructure, starting with the HPC centers. Instead, we advocate a more systematic analysis of the numerical schemes involved in data analyses, which, coupled with software testing, would make the analyses robust to software updates. As a first step, the packages impacting the analyses could be specifically identified and the others updated, which would largely remove vulnerabilities.

Ultimately, software updates should even occur at runtime rather than when the container image is built. Indeed, it is likely that container images used for scientific data analyses be built only occasionally, perhaps every few weeks when a release becomes available, which may not be compatible with the frequency of required security updates. In fact, there is no definite reason for the application software release cycle to be synchronized with security updates, and security updates shouldn't be dependent on application software developers. Instead, we think it would be relevant for analytics engines to (1) systematically apply security updates when containers start, and (2) run software tests provided by application developers, including numerical tests, before running analyses.

Implementing such a workflow, however, requires a long-term endeavour to evaluate broadly the stability of data analysis pipelines, and to develop the associated software tests. For the shorter term, we identified the following recommendations for application developers to reduce the number of security vulnerabilities in container images:

1. *Introduce software dependencies cautiously.* Software dependencies come with a potential security toll that is often neglected. For instance, it can be tempting to add a complete toolbox to implement a relatively minor operation in a data analysis pipeline, such as a data format conversion, while the same functionality might be available in the existing dependencies of the pipeline, albeit in a less convenient way.

2. *Use lightweight base images such as Alpine Linux.* Base images often come with packages that are useful in personal computers or servers, but not in containers dedicated to a specific data analysis. In addition, lightweight distributions define packages with a fine granularity, allowing developers to avoid installing

unnecessary dependencies. For instance, the compressed size of the Alpine Linux base image on DockerHub is only 2.7 MB while that of the Ubuntu base image is 27 MB.

3. *Use OS releases with long-term support.* Security updates are not provided for OS distributions that reached end of life. When a given release of a data analysis pipeline is expected to be used over a long period of time, typically several years as it is common in neurosciences, the life cycle of the distribution release should be considered when choosing a base container image. OS distributions have very different life cycle durations, as long and short life cycles serve different purposes. For instance, among RedHat-based distributions, Fedora release a new version every 6 months and provide maintenance for about a year, while CentOS release every 3-5 years and provide maintenance for 10 years. Similarly, Ubuntu LTS (long-term support) distributions provide free security updates for 5 years, and Debian stable releases are maintained for 3 years.

4. *Install packages, not files.* Vulnerability scanners such as Anchore, Clair or Vuls detect vulnerabilities from the list of packages installed in a container image. Therefore, vulnerabilities contained in software tools installed through direct file download rather than through the package manager would go completely undetected. Domain-specific distributions such as NeuroDebian or NeuroFedora in neuroimaging are useful in this respect.

5. *Minify container images.* The automated minification process that we used in our study is unwieldy for a routine use, as it requires capturing execution traces with ReproZip to reconstruct the graph of package dependencies required for the analysis. In practice, it would be more practical for software developers to identify and remove unnecessary dependencies when they build containers, based on their knowledge of the application.

6. *Run image scanners during continuous integration.* Scanning container images can be a cumbersome process that could be asynchronously executed during continuous integration (CI), through tools such as Travis CI or Circle CI. Including security scans in CI also allows developers to identify vulnerabilities quickly, before new software versions are released.

## 4.2 Contributions

All the collected data are available in our GitHub repository at https://github.com/big-data-lab-team/container-vulnerabilities-paper with a Jupyter notebook to regenerate the figures. We believe that this study is the first of its kind to analyze the container images to see the effect of *image update* and *image minification*. We provide recommendations on how to build container images with a reduced amount of vulnerabilities.

## 4.3 Future Work

Specific attacks that would exploit vulnerabilities in container images against HPC systems should be studied. We believe that such attacks are likely to exist, although attacking HPC systems through containers is challenging due to their relative isolation from the host system. First, under the assumption that legitimate HPC users can be trusted, attackers would have to be remote to the container, either in the same network or on a remote network. Two main types of attacks can be envisaged in these conditions: network-based attacks, exploiting vulnerabilities in network clients installed in the container, and data-based attacks, exploiting vulnerabilities through the processing of malicious data injected through third-party systems.

Several types of escalation attacks could be envisaged once remote attackers gain access to the container, in particular related to (1) using the resources allocated to the container for malicious use, such as storing data in the file system or using CPU cycles, resulting in denial of service for the user running the container and possibly for other HPC users, and (2) attacking a host network service, for instance a scheduler or a file system daemon. Exploits in the host kernel to break out of the container are always possible but unlikely assuming that the host system is maintained by professional system administrators.

# Bibliography


[1] Daniel J Walsh. Bringing new security features to docker. https://opensource.com/business/14/9/security-for-docker. Accessed: 2019-03-04.

[2] Elena Reshetova, Janne Karhunen, Thomas Nyman, and N Asokan. Security of os-level virtualization technologies. In *Nordic Conference on Secure IT Systems*, pages 77–93. Springer, 2014.

[3] David Bernstein. Containers and cloud: From lxc to docker to kubernetes. *IEEE Cloud Computing*, 1(3):81–84, 2014.

[4] containers: real adoption and use cases in 2017. http://i.dell.com/sites/doccontent/business/solutions/whitepapers/en/Documents/Containers_Real_Adoption_2017_Dell_EMC_Forrester_Paper.pdf. Accessed: 2019-02-11.

[5] Charles Anderson. Docker [software engineering]. *IEEE Software*, 32(3):102–c3, 2015.

[6] Set up automated builds. https://docs.docker.com/docker-hub/builds/. Accessed: 2019-02-25.

[7] Antony Martin, Simone Raponi, Théo Combe, and Roberto Di Pietro. Docker ecosystem–vulnerability analysis. *Computer Communications*, 122:30–43, 2018.

[8] Docker hub webhooks. https://docs.docker.com/docker-hub/webhooks/. Accessed: 2019-02-25.

[9] Singularity : a "docker" for hpc environments. https://dev.to/grokcode/

singularity--a-docker-for-hpc-environments-i6p. Accessed: 2019-01-07.

[10] Gregory M Kurtzer, Vanessa Sochat, and Michael W Bauer. Singularity: Scientific containers for mobility of compute. *PloS one*, 12(5):e0177459, 2017.

[11] Gregory M Kurtzer et al. Singularity, 2016.

[12] Justin Cappos, Justin Samuel, Scott Baker, and John H Hartman. A look in the mirror: Attacks on package managers. In *Proceedings of the 15th ACM conference on Computer and communications security*, pages 565–574. ACM, 2008.

[13] Jayanth Gummaraju, Tarun Desikan, and Yoshio Turner. Over 30% of official images in docker hub contain high priority security vulnerabilities. In *Technical Report*. BanyanOps, 2015.

[14] Thanh Bui. Analysis of docker security. *arXiv preprint arXiv:1501.02967*, 2015.

[15] Daniel J Walsh. Are docker containers really secure? https://opensource.com/business/14/7/docker-security-selinux?extIdCarryOver=true&sc_cid=701f2000001OH7EAAW. Accessed: 2019-02-20.

[16] Device whitelist controller. https://www.kernel.org/doc/Documentation/cgroup-v1/devices.txt. Accessed: 2019-02-15.

[17] Andrew Lockhart. *Network Security Hacks*. " O'Reilly Media, Inc.", 2004.

[18] Vivek Ramachandran and Sukumar Nandi. Detecting arp spoofing: An active technique. In *International Conference on Information Systems Security*, pages 239–250. Springer, 2005.

[19] Linux network namespaces. http://www.opencloudblog.com/?p=42.

[20] Docker: Network configuration. https://docs.docker.com/network/. Accessed: 2019-02-02.

[21] Theo Combe, Antony Martin, and Roberto Di Pietro. To docker or not to docker: A security perspective. *IEEE Cloud Computing*, 3(5):54–62, 2016.

[22] Understanding of denial-of-service attacks. https://www.us-cert.gov/ncas/tips/ST04-015. Accessed: 2019-02-24.

[23] Rui Shu, Xiaohui Gu, and William Enck. A study of security vulnerabilities on docker hub. In *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy*, CODASPY '17, pages 269–280, New York, NY, USA, 2017. ACM.

[24] Docker security scanning. https://docs.docker.com/v17.12/docker-cloud/builds/image-scan/. Accessed: 2019-03-04.

[25] Linux capabilities. http://man7.org/linux/man-pages/man7/capabilities.7.html. Accessed: 2019-02-26.

[26] Stephen Smalley, Chris Vance, and Wayne Salamon. Implementing selinux as a linux security module. *NAI Labs Report*, 1(43):139, 2001.

[27] Krzysztof J Gorgolewski, Fidel Alfaro-Almagro, Tibor Auer, Pierre Bellec, Mihai Capotă, M Mallar Chakravarty, Nathan W Churchill, Alexander Li Cohen, R Cameron Craddock, Gabriel A Devenyi, et al. Bids apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS computational biology*, 13(3):e1005209, 2017.

[28] Tristan Glatard, Gregory Kiar, Tristan Aumentado-Armstrong, Natacha Beck, Pierre Bellec, Rémi Bernard, Axel Bonnet, Shawn T Brown, Sorina Camarasu-Pop, Frédéric Cervenansky, et al. Boutiques: a flexible framework to integrate command-line applications in computing platforms. *GigaScience*, 7(5):giy016, 2018.

[29] Rémi Rampin, Fernando Chirigati, Dennis Shasha, Juliana Freire, and Vicky Steeves. Reprozip: The reproducibility packer. *Journal of Open Source Software*, 1(8):107, 2016.

[30] Ed HBM Gronenschild, Petra Habets, Heidi IL Jacobs, Ron Mengelers, Nico Rozendaal, Jim Van Os, and Machteld Marcelis. The effects of freesurfer version, workstation type, and macintosh operating system version on anatomical volume

and cortical thickness measurements. *PloS one*, 7(6):e38234, 2012.

[31] Tristan Glatard, Lindsay B Lewis, Rafael Ferreira da Silva, Reza Adalat, Natacha Beck, Claude Lepage, Pierre Rioux, Marc-Etienne Rousseau, Tarek Sherif, Ewa Deelman, et al. Reproducibility of neuroimaging analyses across operating systems. *Frontiers in neuroinformatics*, 9:12, 2015.