EC601 2019 Fall

Mini-project 2 (Cyber security)

Yi-Wei, Chen
U31239156

Pros and cons on ClusterFuzz and its application

Cyber security is a big topic nowadays especially when everything is going to connect to the internet. As defined by Cisco, Cybersecurity is the practice of protecting systems, networks, and programs from digital attacks.

These cyberattacks are usually aimed at accessing, changing, or destroying sensitive information; extorting money from users; or interrupting normal business processes. A successful cybersecurity approach has multiple layers of protection spread across the computers, networks, programs, or data that one intends to keep safe.

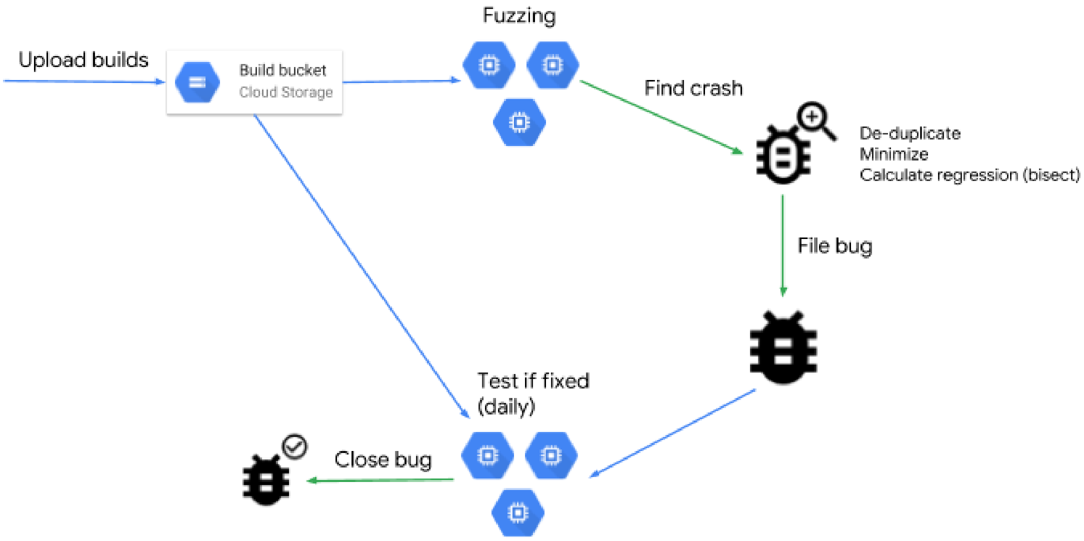
In an organization, the people, processes, and technology must all complement one another to create an effective defense from cyber-attacks. From the people perspective, users must understand and comply with basic data security principles like choosing strong passwords, being wary of attachments in email, and backing up data. From the process perspective, organizations must have a framework for how they deal with both attempted and successful cyber-attacks. One well-respected framework can guide you. It explains how you can identify attacks, protect systems, detect and respond to threats, and recover from successful attacks. From the technology perspective, three main entities must be protected: endpoint devices like computers, smart devices, and routers; networks; and the cloud. Common technology used to protect these entities include next-generation firewalls, DNS filtering, malware protection, antivirus software, and email security solutions.

There are basically four types of attacks in a common situation: Ransomware, Malware, Social engineering and Phishing. Ransomware is a type of malicious software. It is designed to extort money by blocking access to files or the computer system until the ransom is paid. Paying the ransom, however, does not guarantee that the files will come back or the system restored. Malware is a type of software designed to gain unauthorized access or to cause damage to a computer. Social engineering is a tactic that adversaries use to trick you into revealing sensitive information. They can solicit a monetary payment or gain access to your confidential data. Phishing is the practice of sending fraudulent emails that resemble emails from reputable sources. The aim is to steal sensitive data like credit card numbers and login information. It's the most common type of cyber-attack.

After having the fundamental knowledge on what cyber security and cyber-attack are, I am introducing an open source security framework: ClusterFuzz. It is a scalable fuzzing infrastructure that finds security and stability issues in software. Even more, Google uses ClusterFuzz to fuzz the Chrome Browser. Here are some nice features included in ClusterFuzz:

1. Highly scalable. Google's internal instance runs on over 25,000 machines.
2. Accurate deduplication of crashes.
3. Fully automatic bug filing and closing for issue trackers (Monorail only for now).
4. Testcase minimization.
5. Regression finding through bisection.
6. Statistics for analyzing fuzzer performance, and crash rates.
7. Easy-to-use web interface for management and viewing crashes.
8. Firebase authentication.
9. Support for coverage guided fuzzing (e.g. libFuzzer and AFL) and blackbox fuzzing.
10. Learners can get the new updates, versions or suggestions from Github.
11. It is operated in Python which contains lots of library to use.

Here is its terminology:



However, most of the tasks require Google Cloud Platform services. Users can still do some testing locally but only support Linux and Mac.

There are two main components in ClusterFuzz: App Engine instance and a pool of bots. App Engine instance provides a web interface to access crashes, stats and other information. It's also responsible for scheduling regular cron jobs. Bots are machines which run scheduled tasks. They lease tasks from platform specific queues. There are 2 kinds of bots on ClusterFuzz – preemptible and non-preemptible: Preemptible means that the machine can shutdown at any time. On these machines we only run fuzz task. These machines are often cheaper on cloud providers, so it's recommended to scale using these machines.

Non-preemptible machines are not expected to shut down. They are able to run all tasks (including fuzz) and other critical tasks such as progression which must run uninterrupted.

Here are some main tasks bots run:

1. fuzz: Run a fuzzing session.
2. progression: Check if a testcase still reproduces or if it's fixed.
3. regression: Calculate the revision range in which a crash was introduced.
4. minimize: Perform testcase minimization.
5. corpus_pruning: Minimize a corpus to smallest size based on coverage (libFuzzer only).
6. analyze: Run a manually uploaded testcase against a job to see if it crashes.

In General, ClusterFuzz is still a good start for beginners to learn how the security could be done and see some feedback on software developed by themselves. There is also an online forum called CyberPunk for learners to search introductions to a variety of open-source security frameworks and news regarding cyber security issues.

Reference:

1. https://www.cisco.com/c/en/us/products/security/what-is-cybersecurity.html

2. https://n0where.net/

3. https://n0where.net/scalable-fuzzing-infrastructure-clusterfuzz

4. https://google.github.io/clusterfuzz/

Review on my teammate’s report:

After reading Jing Song’s report, I know the operation of Apache Shiro like how we can implement our design based on its framework and its special features compared to other existing open-source security framework. In addition, she also introduces an online community, OWASP, which has plenty of works and projects that others contributed, and I can choose one of them to work on. When working on a same project, I can communicate with other online partner easily to exchange the ideas. This is a resourceful website for me, as a beginner of cyber security, to start with.