

ConPan: A Tool to Analyze Packages in Software Containers

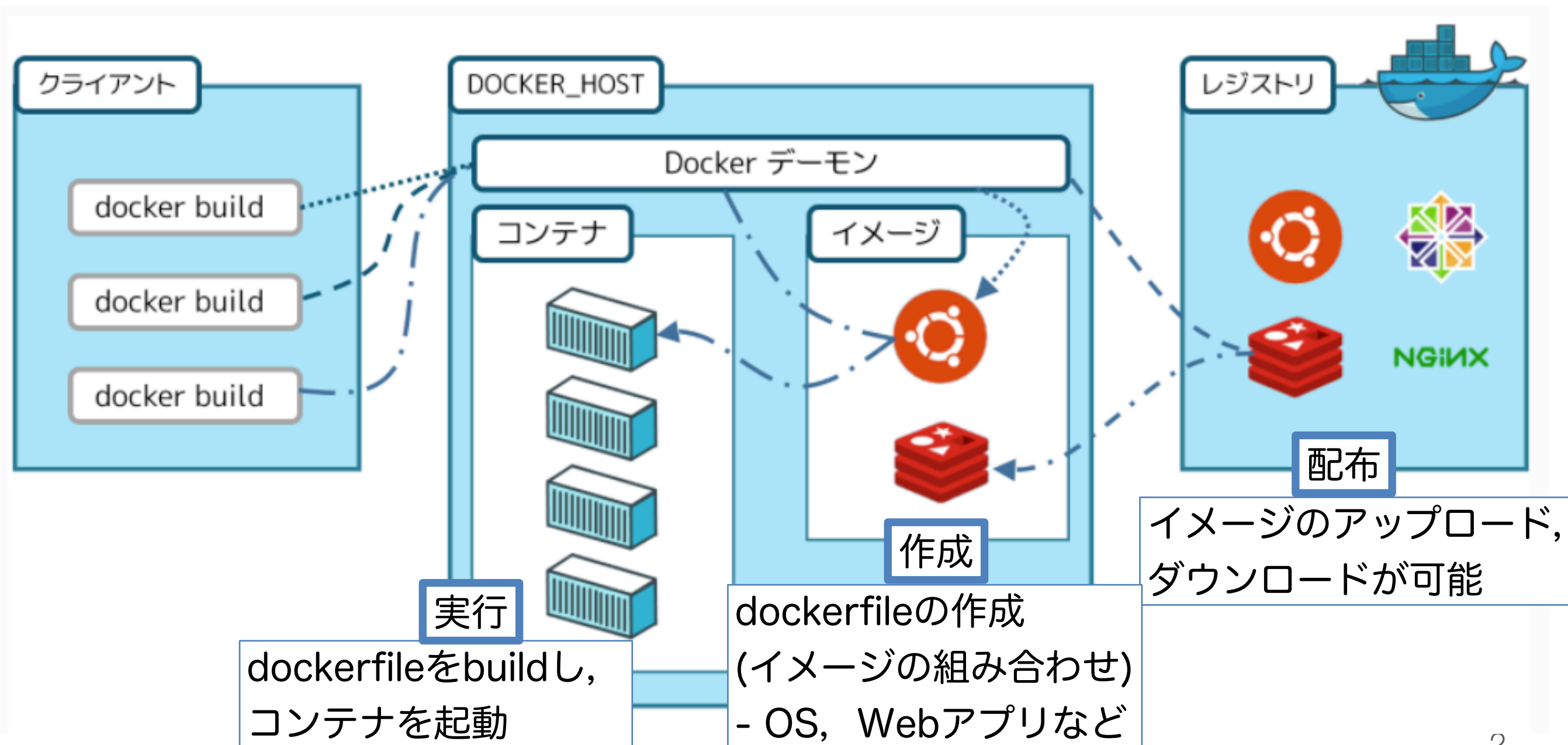
MSR2019

Ahmed Zerouali, Valerio Cosentino, Gregorio Robles,
Jesus M. Gonzalez-Barahona and Tom Mens

作島 大智

Dockerとは

軽量なコンテナ型の仮想環境を作成・配布・実行
するためのプラットフォーム



背景

- Dockerが普及
 - 手軽に同一環境で開発者が作業することが可能
- Dockerコンテナにはバグや脆弱性が含まれている可能性がある
- コンテナイメージを解析するために開発者や研究者は解析ツールを使用する
 - 主にセキュリティ面

解析ツール

- Anchore.com [6]

- Dagda [7]

[6] <https://anchore/>

[7] <https://github.com/eliasgranderubio/dagda>

- Snyk [11]

[11] <https://snyk.io/product/container-vulnerability-management/>

問題点

セキュリティ脆弱性以外を調査する場合,
(コンテナのバージョン遅れや, 品質問題の評価)
自分でスクリプトを用意する必要がある

目的

Dockerコンテナにインストールされたソフトウェアパッケージのデータを収集して取得することを簡素化する

ConPanツールの提案

コンテナ内のパッケージに関するデータを複数のソースから簡単かつ一貫性のある方法で追跡・取得

ConPan : <https://github.com/neglectos/ConPan>

ConPanの概要

Dockerイメージをpullして、実行

インストールされているパッケージを特定

パッケージマネージャを特定

既知の脆弱性報告や、その他報告されている
バグや品質問題を検索

1つのJSONファイルと4つのpandas DataFrameに出力



ConPanの概要

Dockerイメージをpullして、実行

インストールされているパッケージを特定

パッケージマネージャを特定

パッケージがいつどこでリリースされたかという
過去の情報を含むデータセットに依存
ex) Debian パッケージに対応するデータセットは
Debian アーカイブ [20]

ConPanの概要

脆弱性：セキュリティトラッカー

バグ：バグトラッカー

ex) Debian の場合、公式のセキュリティトラッカー [21],
Ultimate Debian Database [22]

既知の脆弱性報告や、その他報告されている
バグや品質問題を検索する

パッケージの最新版でも修正されていない or イメージ内
バージョン以降で修正されている場合、マークをつける

出力結果の例

Results:

General information about the Docker image: google/
mysql

- description: MySQL server for Google
Compute Engine
- star_count: 18
- pull_count: 46692
- full_size: 96687899
- last_updated: 2015-11-13T01:19:18.235940

一般的な情報
(JSONファイル)

Results about installed packages in: google/mysql

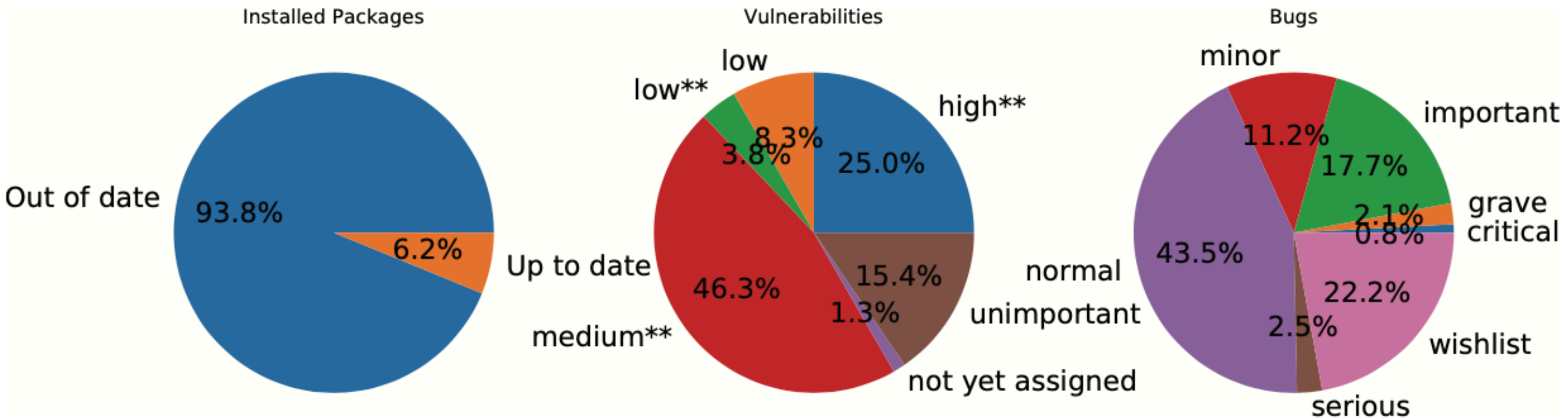
- # installed packages: 144
- # tracked packages: 81
- # vulnerabilities: 240
- # bugs: 474

パッケージ情報
(DataFrame型)

出力結果の利用

- matplotlib や seaborn ライブラリを使って可視化
- データフレームを JSON ファイルに変換して
NoSQLドキュメントベースのストレージに保存
- CSV 形式の生データとして保存

出力結果を可視化



- 古いパッケージの割合が高い (93.8%)
- 脆弱性240件, バグ474件と多い
- 深刻度は中～高程度

まとめ

- Docker Hubイメージとそのインストールされたパッケージの分析を簡素化するツールを開発
 - イメージのパッケージが含む脆弱性・バグ検出
 - 可視化しやすいデータ形式

所感

- 複数のデータソースの情報をまとめてくれるツールなので便利だと感じた.
- 提案手法で収集したデータ(特に脆弱性・バグ)は研究に活用できるのだろうかと疑問に思った.
- Debianが例にあるが, 他のOSパッケージでも同様のことができるのか気になった.