Should you Upgrade Official Docker Hub Images in Production Environments?

Sara Gholami
Hamzeh Khazaei
Cor-Paul Bezemer
Containers are standalone, lightweight units of software that package software code and all of its dependencies.
DockerHub is Docker’s default public registry

130 B
Total pulls of images

3.5 M
Docker Hub Repositories

5 M
Docker Hub Users
Docker images facilitate the deployment of applications in production environments.
A Docker image is created from a DockerFile

```
FROM ubuntu  # Adds packages required by ubuntu

RUN apt-get install -y software-properties-common python  # Adds python packages
RUN add-apt-repository ppa:chris-lea/node.js
RUN apt-get update
RUN apt-get install -y nodejs

RUN mkdir /var/www
ADD app.js /var/www/app.js

CMD ["/usr/bin/node", "/var/www/app.js"]
```

Adds node packages
Using Docker images in production environments is easy, but it increases the risk of package changes

We quantify the package changes in Docker images across different applications
Our methodology to study package changes in Docker images

- Selecting Repositories
- Collecting Image Tags
- Collecting Packages and Latest Update Dates
- Identifying Package Changes
Our methodology to study package changes in Docker images

Selecting Repositories
- Docker Hub
  - Selecting repositories
    - 160 official repositories

Collecting Image Tags
- Docker Hub
  - Retrieving repository tags
  - 37K repository tags

Collecting Packages and Latest Update Dates
- Docker Hub
  - Pull 37K images
  - Extract packages
  - Get 37K latest update dates
  - List of packages and latest update dates

Identifying Package Changes
- Break down repositories into branches
- Sort images in branches
- Identify package changes
- Package changes

Which types of applications tend to have more package changes?
There are four types of repositories on Docker Hub

**Community**
(Developed by community developers)
3.3M 99.9%

**Verified**
(Developed by third-party software vendors)
339 < 1%

**Official**
(Reviewed by the Docker team)
160 < 1%

**Certified**
(Special category of verified repositories)
51 < 1%
Docker official repositories are more popular than other types of repositories
Docker official repositories are downloaded more often than other types of repositories
Our methodology to study package changes in Docker images

Selecting Repositories
- Docker Hub
- Selecting repositories
- 160 official repositories

Collecting Image Tags
- Docker Hub
- Retrieving repository tags
- 37K repository tags

Collecting Packages and Latest Update Dates
- Docker Hub
- Pull 37K images
- Extract packages
- Get 37K latest update dates
- List of packages and latest update dates

Identifying Package Changes
- Break down repositories into branches
A repository can have several branches which should be analyzed separately
Our methodology to study package changes in Docker images

Selecting Repositories
- Docker Hub
- Selecting repositories
- 160 official repositories

Collecting Image Tags
- Docker Hub
- Retrieving repository tags
- 37K repository tags

Collecting Packages and Latest Update Dates
- Docker Hub
- Pull 37K images
- Extract packages
- Get 37K latest update dates
- List of packages and latest update dates

Identifying Package Changes
- Break down repositories into branches
- Sort images in branches
- Identify package changes
There are three types of version changes: major, minor, and patch.
Our methodology to study package changes in Docker images

- **Selecting Repositories**
  - Docker Hub
  - Selecting repositories
  - 160 official repositories

- **Collecting Image Tags**
  - Docker Hub
  - Retrieving repository tags
  - 37K repository tags

- **Collecting Packages and Latest Update Dates**
  - Docker Hub
  - Pull 37K images
  - Extract packages
  - Get 37K latest update dates
  - List of packages and latest update dates

- **Identifying Package Changes**
  - Break down repositories into branches
  - Sort images in branches
  - Identify package changes
  - Package changes

Which types of applications tend to have more package changes?
Results and Discussions
There is a median of 8.6 upgrades per image across official Docker images.
There is a median of 2.1 downgrades per image across official Docker images.
Images of *Analytics* applications are the least stable
The packages that are changed the most often are common utility packages

From over 9K packages used in official Docker images, the top ones with the highest number of changes are common utility packages such as *tzdata and base-files*
In-place upgrading Docker images in production environments has risk of package changes

Practitioners need to be cautious when doing in-place upgrades of images from the official Docker Hub repositories as in all studied applications, many packages are changing.
Our methodology to study package changes in Docker images

Selecting Repositories
- Docker Hub
  - Selecting repositories
    - 160 official repositories

Collecting Image Tags
- Docker Hub
  - Retrieving repository tags
    - 37K repository tags

Collecting Packages and Latest Update Dates
- Docker Hub
  - Pull 37K images
    - Extract packages
      - List of packages and latest update dates
  - Get 37K latest update dates

Identifying Package Changes
- Break down repositories into branches
  - Sort images in branches
    - Identify package changes
      - Package changes

Which types of applications tend to have more package changes?
Our methodology to study package changes in Docker images
There is a median of 8.6 upgrades per image across official Docker images.
Our methodology to study package changes in Docker images

There is a median of 8.6 upgrades per image across official Docker images
There is a median of 2.1 downgrades per image across official Docker images
Our methodology to study package changes in Docker images

- Selecting Repositories
- Collecting Image Tags
- Collecting Packages and Latest Update Dates
- Identifying Package Changes

There is a median of 2.1 downgrades per image across official Docker images

There is a median of 8.6 upgrades per image across official Docker images

Sara Gholami
sgholami@ualberta.ca
https://asgaard.ece.ualberta.ca/