**Learning Docker in the Play with Docker classroom.**

Learning some basic Docker commands and a simple build-ship-run workflow, starting by running some simple containers, then used a Dockerfile to build a custom app. Finally, used bind mounts to modify a running container.

# WARNING!!!! #

# This is a sandbox environment. Using personal credentials #

# is HIGHLY! discouraged. Any consequences of doing so are #

# completely the user's responsibilites. #

# #

# The PWD team. #

###############################################################

[node1] (local) root@192.168.0.8 ~

$ git clone https://github.com/dockersamples/linux\_tweet\_app

Cloning into 'linux\_tweet\_app'...

remote: Enumerating objects: 14, done.

remote: Total 14 (delta 0), reused 0 (delta 0), pack-reused 14

Receiving objects: 100% (14/14), 10.76 KiB | 10.76 MiB/s, done.

Resolving deltas: 100% (5/5), done.

[node1] (local) root@192.168.0.8 ~

$ docker container run alpine hostname

Unable to find image 'alpine:latest' locally

latest: Pulling from library/alpine

59bf1c3509f3: Pull complete

Digest: sha256:21a3deaa0d32a8057914f36584b5288d2e5ecc984380bc0118285c70fa8c9300

Status: Downloaded newer image for alpine:latest

1d801b37f3ac

[node1] (local) root@192.168.0.8 ~

$ docker container ls --all

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

1d801b37f3ac alpine "hostname" 3 seconds ago Exited (0) 2 seconds ago heuristic\_joliot

[node1] (local) root@192.168.0.8 ~

$ docker container run --interactive --tty --rm ubuntu bash

Unable to find image 'ubuntu:latest' locally

latest: Pulling from library/ubuntu

ea362f368469: Pull complete

Digest: sha256:b5a61709a9a44284d88fb12e5c48db0409cfad5b69d4ff8224077c57302df9cf

Status: Downloaded newer image for ubuntu:latest

root@139f0ff609cb:/# ls /

bin dev home lib32 libx32 mnt proc run srv tmp var

boot etc lib lib64 media opt root sbin sys usr

root@139f0ff609cb:/# ps aux

USER PID %CPU %MEM VSZ RSS TTY STAT START TIME COMMAND

root 1 1.0 0.0 4108 3568 pts/0 Ss 04:18 0:00 bash

root 11 0.0 0.0 5896 2924 pts/0 R+ 04:18 0:00 ps aux

root@139f0ff609cb:/# cat /etc/issue

Ubuntu 20.04.3 LTS \n \l

root@139f0ff609cb:/# exit

exit

[node1] (local) root@192.168.0.8 ~

$ cat /etc/issue

Welcome to Alpine Linux 3.12

Kernel \r on an \m (\l)

[node1] (local) root@192.168.0.8 ~

$ docker container run \

> --detach \

> --name mydb \

> -e MYSQL\_ROOT\_PASSWORD=my-secret-pw \

> mysql:latest

Unable to find image 'mysql:latest' locally

latest: Pulling from library/mysql

72a69066d2fe: Pull complete

93619dbc5b36: Pull complete

99da31dd6142: Pull complete

626033c43d70: Pull complete

37d5d7efb64e: Pull complete

ac563158d721: Pull complete

d2ba16033dad: Pull complete

688ba7d5c01a: Pull complete

00e060b6d11d: Pull complete

1c04857f594f: Pull complete

4d7cfa90e6ea: Pull complete

e0431212d27d: Pull complete

Digest: sha256:e9027fe4d91c0153429607251656806cc784e914937271037f7738bd5b8e7709

Status: Downloaded newer image for mysql:latest

1be7e404cc347e10a9d0eb3f0fe3e47463a31cbd2ecbfe2dcfe7e0dbbc3c0ec8

[node1] (local) root@192.168.0.8 ~

$ docker container ls

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

1be7e404cc34 mysql:latest "docker-entrypoint.s…" 3 seconds ago Up 2 seconds 3306/tcp, 33060/tcp mydb

[node1] (local) root@192.168.0.8 ~

$ docker container logs mydb

2022-01-08 04:18:25+00:00 [Note] [Entrypoint]: Entrypoint script for MySQL Server 8.0.27-1debian10 started.

2022-01-08 04:18:26+00:00 [Note] [Entrypoint]: Switching to dedicated user 'mysql'

2022-01-08 04:18:26+00:00 [Note] [Entrypoint]: Entrypoint script for MySQL Server 8.0.27-1debian10 started.

2022-01-08 04:18:26+00:00 [Note] [Entrypoint]: Initializing database files

2022-01-08T04:18:26.162554Z 0 [System] [MY-013169] [Server] /usr/sbin/mysqld (mysqld 8.0.27)initializing of server in progress as process 44

2022-01-08T04:18:26.173643Z 1 [System] [MY-013576] [InnoDB] InnoDB initialization has started.

2022-01-08T04:18:26.522145Z 1 [System] [MY-013577] [InnoDB] InnoDB initialization has ended.

2022-01-08T04:18:27.817124Z 0 [Warning] [MY-013746] [Server] A deprecated TLS version TLSv1 is enabled for channel mysql\_main

2022-01-08T04:18:27.817169Z 0 [Warning] [MY-013746] [Server] A deprecated TLS version TLSv1.1 is enabled for channel mysql\_main

2022-01-08T04:18:27.841876Z 6 [Warning] [MY-010453] [Server] root@localhost is created with an empty password ! Please consider switching off the --initialize-insecure option.

[node1] (local) root@192.168.0.8 ~

$ docker container top mydb

PID USER TIME COMMAND

970 999 0:00 {docker-entrypoi} /bin/bash /usr/local/bin/docker-entrypoint.sh mysqld

1173 999 0:03 mysqld --daemonize --skip-networking --default-time-zone=SYSTEM --socket=/var/run/mysqld/mysqld.sock

1221 999 0:00 mysql\_tzinfo\_to\_sql /usr/share/zoneinfo

1222 999 0:00 sed s/Local time zone must be set--see zic manual page/FCTY/

1223 999 0:00 {docker-entrypoi} /bin/bash /usr/local/bin/docker-entrypoint.sh mysqld

1225 999 0:00 mysql --defaults-extra-file=/dev/fd/63 --protocol=socket -uroot -hlocalhost --socket=/var/run/mysqld/mysqld.sock --comments --database=mysql

[node1] (local) root@192.168.0.8 ~

$ docker exec -it mydb \

> mysql --user=root --password=$MYSQL\_ROOT\_PASSWORD --version

mysql: [Warning] Using a password on the command line interface can be insecure.

mysql Ver 8.0.27 for Linux on x86\_64 (MySQL Community Server - GPL)

[node1] (local) root@192.168.0.8 ~

$ docker exec -it mydb sh

# mysql --user=root --password=$MYSQL\_ROOT\_PASSWORD --version

mysql: [Warning] Using a password on the command line interface can be insecure.

mysql Ver 8.0.27 for Linux on x86\_64 (MySQL Community Server - GPL)

# exit

[node1] (local) root@192.168.0.8 ~

$ cd ~/linux\_tweet\_app

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ cat Dockerfile

FROM nginx:latest

COPY index.html /usr/share/nginx/html

COPY linux.png /usr/share/nginx/html

EXPOSE 80 443

CMD ["nginx", "-g", "daemon off;"]

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ echo $DOCKERID

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ export DOCKERID=paru97

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ echo paru97

paru97

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image build --tag paru97/linux\_tweet\_app:1.0 .

Sending build context to Docker daemon 92.16kB

Step 1/5 : FROM nginx:latest

latest: Pulling from library/nginx

a2abf6c4d29d: Pull complete

a9edb18cadd1: Pull complete

589b7251471a: Pull complete

186b1aaa4aa6: Pull complete

b4df32aa5a72: Pull complete

a0bcbecc962e: Pull complete

Digest: sha256:0d17b565c37bcbd895e9d92315a05c1c3c9a29f762b011a10c54a66cd53c9b31

Status: Downloaded newer image for nginx:latest

---> 605c77e624dd

Step 2/5 : COPY index.html /usr/share/nginx/html

---> 678793c0c77b

Step 3/5 : COPY linux.png /usr/share/nginx/html

---> 8a763cb2d8a2

Step 4/5 : EXPOSE 80 443

---> Running in 6f22818b5b2e

Removing intermediate container 6f22818b5b2e

---> 9940ceb8bef0

Step 5/5 : CMD ["nginx", "-g", "daemon off;"]

---> Running in 445891a78ead

Removing intermediate container 445891a78ead

---> a5f8b7dbd4a6

Successfully built a5f8b7dbd4a6

Successfully tagged paru97/linux\_tweet\_app:1.0

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container run \

> --detach \

> --publish 80:80 \

> --name linux\_tweet\_app \

> paru97/linux\_tweet\_app:1.0

adaffc24f975263928afa05fb1f0cd0a2f85e013b09089213d0a5289191327a6

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container rm --force linux\_tweet\_app

linux\_tweet\_app

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container run \

> --detach \

> --publish 80:80 \

> --name linux\_tweet\_app \

> --mount type=bind,source="$(pwd)",target=/usr/share/nginx/html \

> paru97/linux\_tweet\_app:1.0

06ad9f91f6bcf38e400bfc41ba462ea75604bdf78dda9f3b3b5ee80c6dc5dd1b

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ cp index-new.html index.html

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker rm --force linux\_tweet\_app

linux\_tweet\_app

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container run \

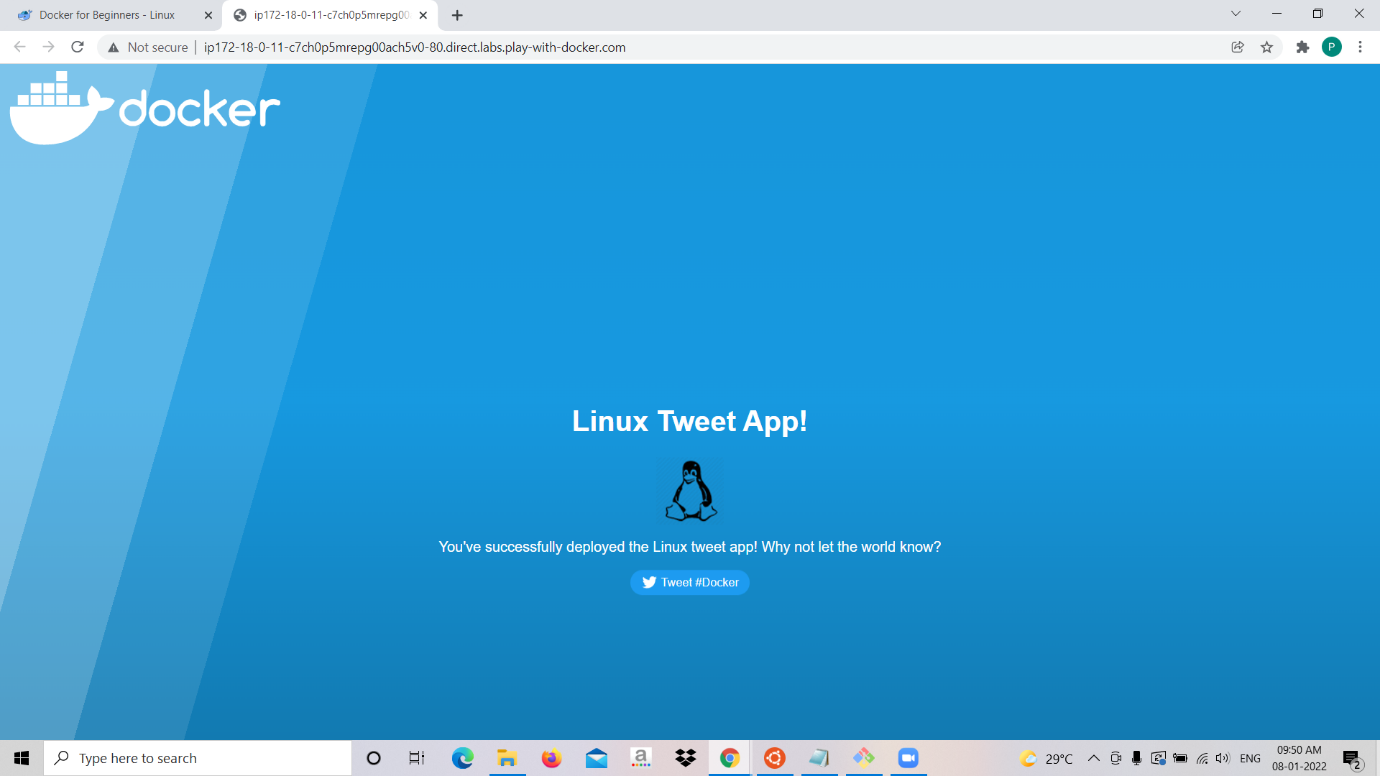
> --detach \

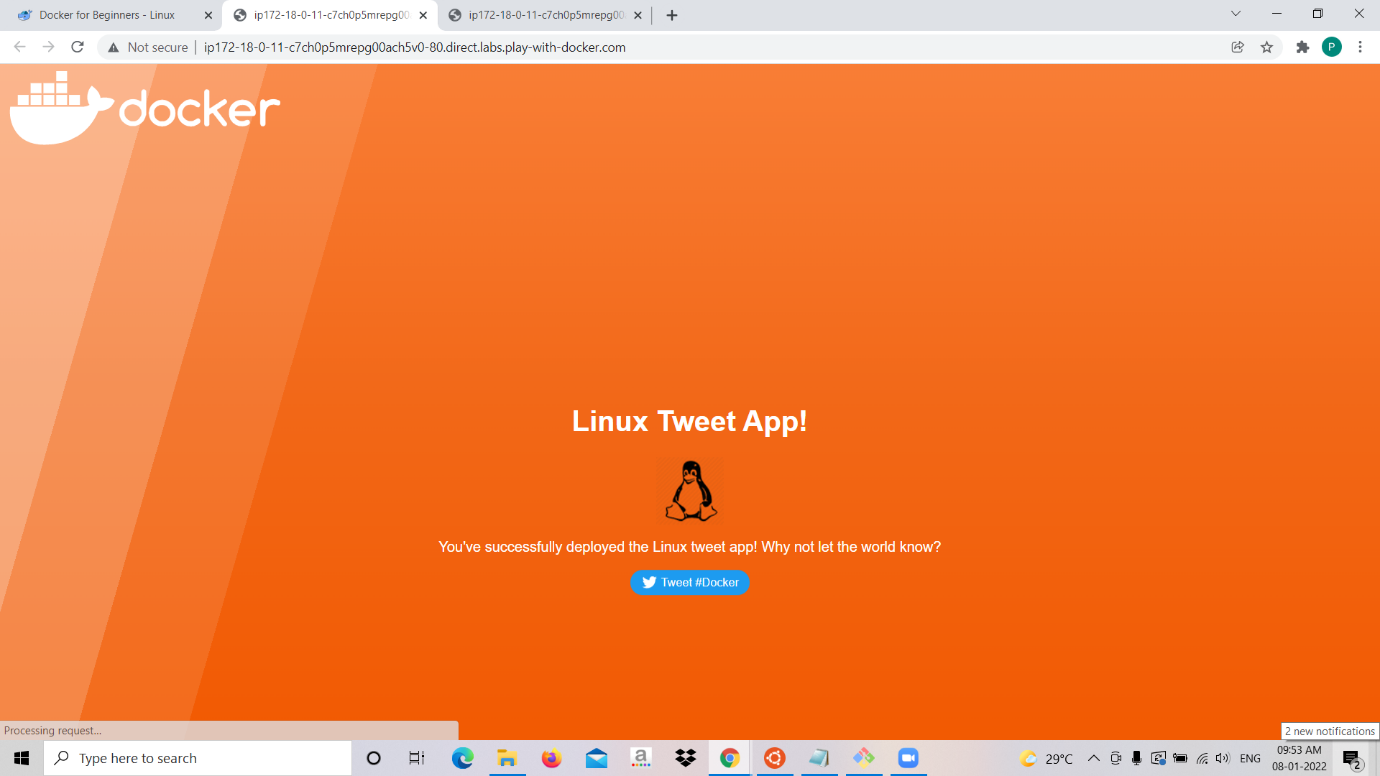
> --publish 80:80 \

> --name linux\_tweet\_app \

> paru97/linux\_tweet\_app:1.0

7669e40132a13d02dfac55439b65357bf0ab278531ea3d2fdd4360896f5e846e





[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image build --tag paru97/linux\_tweet\_app:2.0 .

Sending build context to Docker daemon 92.16kB

Step 1/5 : FROM nginx:latest

---> 605c77e624dd

Step 2/5 : COPY index.html /usr/share/nginx/html

---> e96e6cb37624

Step 3/5 : COPY linux.png /usr/share/nginx/html

---> b30f001495f1

Step 4/5 : EXPOSE 80 443

---> Running in 58114abe1d7a

Removing intermediate container 58114abe1d7a

---> 3100ff60c8b8

Step 5/5 : CMD ["nginx", "-g", "daemon off;"]

---> Running in a8a8f3669c1e

Removing intermediate container a8a8f3669c1e

---> 1c382bb66ca9

Successfully built 1c382bb66ca9

Successfully tagged paru97/linux\_tweet\_app:2.0

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image ls

REPOSITORY TAG IMAGE ID CREATED SIZE

paru97/linux\_tweet\_app 2.0 1c382bb66ca9 4 seconds ago 142MB

paru97/linux\_tweet\_app 1.0 a5f8b7dbd4a6 5 minutes ago 142MB

ubuntu latest d13c942271d6 26 hours ago 72.8MB

nginx latest 605c77e624dd 9 days ago 141MB

mysql latest 3218b38490ce 2 weeks ago 516MB

alpine latest c059bfaa849c 6 weeks ago 5.59MB

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container run \

> --detach \

> --publish 80:80 \

> --name linux\_tweet\_app \

> paru97/linux\_tweet\_app:2.0

f310d639fb31c85ffc7b3593c46f140b3a212f744ba6dd7314b74a684b6f3728

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker container run --detach --publish 80:80 --name linux\_tweet\_app paru97/linux\_tweet\_app:1.0

docker: Error response from daemon: Conflict. The container name "/linux\_tweet\_app" is already in use by container "f310d639fb31c85ffc7b3593c46f140b3a212f744ba6dd7314b74a684b6f3728". You have to remove (or rename) that container to be able to reuse that name.

See 'docker run --help'.

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image ls -f reference="paru97/\*"

REPOSITORY TAG IMAGE ID CREATED SIZE

paru97/linux\_tweet\_app 2.0 1c382bb66ca9 2 minutes ago 142MB

paru97/linux\_tweet\_app 1.0 a5f8b7dbd4a6 7 minutes ago 142MB

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker login

Login with your Docker ID to push and pull images from Docker Hub. If you don't have a Docker ID, head over to https://hub.docker.com to create one.

Username: paru97

Password:

WARNING! Your password will be stored unencrypted in /root/.docker/config.json.

Configure a credential helper to remove this warning. See

https://docs.docker.com/engine/reference/commandline/login/#credentials-store

Login Succeeded

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image push paru97/linux\_tweet\_app:1.0

The push refers to repository [docker.io/paru97/linux\_tweet\_app]

0d1256717388: Pushed

bfdf06b20db0: Pushed

d874fd2bc83b: Mounted from library/nginx

32ce5f6a5106: Mounted from library/nginx

f1db227348d0: Mounted from library/nginx

b8d6e692a25e: Mounted from library/nginx

e379e8aedd4d: Mounted from library/nginx

2edcec3590a4: Layer already exists

1.0: digest: sha256:85be2a40857923e958900c36f7de2767818516655c4cb254f054aa008bffd4a1 size: 1985

[node1] (local) root@192.168.0.8 ~/linux\_tweet\_app

$ docker image push paru97/linux\_tweet\_app:2.0

The push refers to repository [docker.io/paru97/linux\_tweet\_app]

625a25e5c68e: Pushed

b1668ec47465: Pushed

d874fd2bc83b: Layer already exists

32ce5f6a5106: Layer already exists

f1db227348d0: Layer already exists

b8d6e692a25e: Layer already exists

e379e8aedd4d: Layer already exists

2edcec3590a4: Layer already exists

2.0: digest: sha256:e7f0ebcc481b5b44333d1937e5ffba0e4321c1954aee061260acdbd2457011f6 size: 1985

Application Containerization and Microservice Orchestration

learning about basic application containerization using Docker and running various components of an application as microservices. We will utilize [Docker Compose](https://docs.docker.com/compose/) for orchestration during the development.

Starting from a basic Python script that scrapes links from a given web page and gradually evolve it into a multi-service application stack. The demo code is available in the [Link Extractor](https://github.com/ibnesayeed/linkextractor) repo. The code is organized in steps that incrementally introduce changes and new concepts. After completion, the application stack will contain the following microservices:

* A web application written in PHP and served using Apache that takes a URL as the input and summarizes extracted links from it
* The web application talks to an API server written in Python (and Ruby) that takes care of the link extraction and returns a JSON response
* A Redis cache that is used by the API server to avoid repeated fetch and link extraction for pages that are already scraped

[node1] (local) root@192.168.0.28 ~

$ git clone https://github.com/ibnesayeed/linkextractor.git

Cloning into 'linkextractor'...

remote: Enumerating objects: 144, done.

remote: Counting objects: 100% (4/4), done.

remote: Compressing objects: 100% (4/4), done.

remote: Total 144 (delta 0), reused 1 (delta 0), pack-reused 140Receiving objects: 100% (144/144), 44.55 KiB | 7.42 MiB/s, done.Resolving deltas: 100% (43/43), done.[node1] (local) root@192.168.0.28 ~$ cd linkextractor[node1] (local) root@192.168.0.28 ~/linkextractor$ git checkout demoBranch 'demo' set up to track remote branch 'demo' from 'origin'.Switched to a new branch 'demo'[node1] (local) root@192.168.0.28 ~/linkextractor$ git checkout step0Branch 'step0' set up to track remote branch 'step0' from 'origin'.Switched to a new branch 'step0'[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── README.md

└── linkextractor.py

0 directories, 2 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat linkextractor.py

#!/usr/bin/env python

import sys

import requests

from bs4 import BeautifulSoup

res = requests.get(sys.argv[-1])

soup = BeautifulSoup(res.text, "html.parser")

for link in soup.find\_all("a"):

print(link.get("href"))

[node1] (local) root@192.168.0.28 ~/linkextractor

$ ./linkextractor.py http://example.com/

bash: ./linkextractor.py: Permission denied

[node1] (local) root@192.168.0.28 ~/linkextractor

$ ls -l linkextractor.py

-rw-r--r-- 1 root root 220 Jan 8 04:44 linkextractor.py

[node1] (local) root@192.168.0.28 ~/linkextractor

$ python3 linkextractor.py

Traceback (most recent call last):

File "linkextractor.py", line 5, in <module>

from bs4 import BeautifulSoup

ModuleNotFoundError: No module named 'bs4'

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step1

Branch 'step1' set up to track remote branch 'step1' from 'origin'.

Switched to a new branch 'step1'

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── Dockerfile

├── README.md

└── linkextractor.py

0 directories, 3 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat Dockerfile

FROM python:3

LABEL maintainer="Sawood Alam <@ibnesayeed>"

RUN pip install beautifulsoup4

RUN pip install requests

WORKDIR /app

COPY linkextractor.py /app/

RUN chmod a+x linkextractor.py

ENTRYPOINT ["./linkextractor.py"]

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker image build -t linkextractor:step1 .

Sending build context to Docker daemon 107.5kB

Step 1/8 : FROM python:3

3: Pulling from library/python

0e29546d541c: Pull complete

9b829c73b52b: Pull complete

cb5b7ae36172: Pull complete

6494e4811622: Pull complete

6f9f74896dfa: Pull complete

fcb6d5f7c986: Pull complete

290438add9da: Pull complete

ab11df61f44a: Pull complete

de4793a5fa46: Pull complete

Digest: sha256:dbbfcbf95f6b596d2be1d8f3b368016619f78f829facf6f2e361bea1151794e5

Status: Downloaded newer image for python:3

---> a5d7930b60cc

Step 2/8 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Running in afbec33ed849

Removing intermediate container afbec33ed849

---> 7d018a9f0c68

Step 3/8 : RUN pip install beautifulsoup4

---> Running in 6a03f0c0daa4

Collecting beautifulsoup4

Downloading beautifulsoup4-4.10.0-py3-none-any.whl (97 kB)

Collecting soupsieve>1.2

Downloading soupsieve-2.3.1-py3-none-any.whl (37 kB)

Installing collected packages: soupsieve, beautifulsoup4

Successfully installed beautifulsoup4-4.10.0 soupsieve-2.3.1

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv

WARNING: You are using pip version 21.2.4; however, version 21.3.1 is available.

You should consider upgrading via the '/usr/local/bin/python -m pip install --upgrade pip' command.

Removing intermediate container 6a03f0c0daa4

---> 80fb985c973c

Step 4/8 : RUN pip install requests

---> Running in fd956eba99ce

Collecting requests

Downloading requests-2.27.1-py2.py3-none-any.whl (63 kB)

Collecting charset-normalizer~=2.0.0

Downloading charset\_normalizer-2.0.10-py3-none-any.whl (39 kB)

Collecting idna<4,>=2.5

Downloading idna-3.3-py3-none-any.whl (61 kB)

Collecting urllib3<1.27,>=1.21.1

Downloading urllib3-1.26.8-py2.py3-none-any.whl (138 kB)

Collecting certifi>=2017.4.17

Downloading certifi-2021.10.8-py2.py3-none-any.whl (149 kB)

Installing collected packages: urllib3, idna, charset-normalizer, certifi, requests

Successfully installed certifi-2021.10.8 charset-normalizer-2.0.10 idna-3.3 requests-2.27.1 urllib3-1.26.8

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv

WARNING: You are using pip version 21.2.4; however, version 21.3.1 is available.

You should consider upgrading via the '/usr/local/bin/python -m pip install --upgrade pip' command.

Removing intermediate container fd956eba99ce

---> 9c01cf0ada9f

Step 5/8 : WORKDIR /app

---> Running in ad2a2a0dcd6e

Removing intermediate container ad2a2a0dcd6e

---> 96641ad3465d

Step 6/8 : COPY linkextractor.py /app/

---> 8e92cc3d9609

Step 7/8 : RUN chmod a+x linkextractor.py

---> Running in d8c66c1a924c

Removing intermediate container d8c66c1a924c

---> deb83d0c6965

Step 8/8 : ENTRYPOINT ["./linkextractor.py"]

---> Running in 2413a373b067

Removing intermediate container 2413a373b067

---> 995226033a1f

Successfully built 995226033a1f

Successfully tagged linkextractor:step1

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker image ls

REPOSITORY TAG IMAGE ID CREATED SIZE

linkextractor step1 995226033a1f 2 seconds ago 927MB

python 3 a5d7930b60cc 2 weeks ago 917MB

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container run -it --rm linkextractor:step1 http://example.com/

https://www.iana.org/domains/example

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container run -it --rm linkextractor:step1 https://training.play-with-docker.com/

/

/about/

#ops

#dev

/ops-stage1

/ops-stage2

/ops-stage3

/dev-stage1

/dev-stage2

/dev-stage3

/alacart

https://twitter.com/intent/tweet?text=Play with Docker Classroom&url=https://training.play-with-docker.com/&via=docker&related=docker

https://facebook.com/sharer.php?u=https://training.play-with-docker.com/

https://plus.google.com/share?url=https://training.play-with-docker.com/

http://www.linkedin.com/shareArticle?mini=true&url=https://training.play-with-docker.com/&title=Play%20with%20Docker%20Classroom&source=https://training.play-with-docker.com

https://www.docker.com/dockercon/

https://www.docker.com/dockercon/

https://dockr.ly/slack

https://www.docker.com/legal/docker-terms-service

https://www.docker.com

https://www.facebook.com/docker.run

https://twitter.com/docker

https://www.github.com/play-with-docker/play-with-docker.github.io

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step2

Branch 'step2' set up to track remote branch 'step2' from 'origin'.

Switched to a new branch 'step2'

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── Dockerfile

├── README.md

└── linkextractor.py

0 directories, 3 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat linkextractor.py

#!/usr/bin/env python

import sys

import requests

from bs4 import BeautifulSoup

from urllib.parse import urljoin

def extract\_links(url):

res = requests.get(url)

soup = BeautifulSoup(res.text, "html.parser")

base = url

# TODO: Update base if a <base> element is present with the href attribute

links = []

for link in soup.find\_all("a"):

links.append({

"text": " ".join(link.text.split()) or "[IMG]",

"href": urljoin(base, link.get("href"))

})

return links

if \_\_name\_\_ == "\_\_main\_\_":

if len(sys.argv) != 2:

print("\nUsage:\n\t{} <URL>\n".format(sys.argv[0]))

sys.exit(1)

for link in extract\_links(sys.argv[-1]):

print("[{}]({})".format(link["text"], link["href"]))

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker image build -t linkextractor:step2 .

Sending build context to Docker daemon 110.1kB

Step 1/8 : FROM python:3

---> a5d7930b60cc

Step 2/8 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 7d018a9f0c68

Step 3/8 : RUN pip install beautifulsoup4

---> Using cache

---> 80fb985c973c

Step 4/8 : RUN pip install requests

---> Using cache

---> 9c01cf0ada9f

Step 5/8 : WORKDIR /app

---> Using cache

---> 96641ad3465d

Step 6/8 : COPY linkextractor.py /app/

---> 776358119bd0

Step 7/8 : RUN chmod a+x linkextractor.py

---> Running in 63b666bde5fe

Removing intermediate container 63b666bde5fe

---> e36ba3bac2c2

Step 8/8 : ENTRYPOINT ["./linkextractor.py"]

---> Running in 6b04b8af2d07

Removing intermediate container 6b04b8af2d07

---> 82c626513e57

Successfully built 82c626513e57

Successfully tagged linkextractor:step2

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker image ls

REPOSITORY TAG IMAGE ID CREATED SIZE

linkextractor step2 82c626513e57 1 second ago 927MB

linkextractor step1 995226033a1f About a minute ago 927MB

python 3 a5d7930b60cc 2 weeks ago 917MB

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container run -it --rm linkextractor:step2 https://training.play-with-docker.com/

[Play with Docker classroom](https://training.play-with-docker.com/)

[About](https://training.play-with-docker.com/about/)

[IT Pros and System Administrators](https://training.play-with-docker.com/#ops)

[Developers](https://training.play-with-docker.com/#dev)

[Stage 1: The Basics](https://training.play-with-docker.com/ops-stage1)

[Stage 2: Digging Deeper](https://training.play-with-docker.com/ops-stage2)

[Stage 3: Moving to Production](https://training.play-with-docker.com/ops-stage3)

[Stage 1: The Basics](https://training.play-with-docker.com/dev-stage1)

[Stage 2: Digging Deeper](https://training.play-with-docker.com/dev-stage2)

[Stage 3: Moving to Staging](https://training.play-with-docker.com/dev-stage3)

[Full list of individual labs](https://training.play-with-docker.com/alacart)

[[IMG]](https://twitter.com/intent/tweet?text=Play with Docker Classroom&url=https://training.play-with-docker.com/&via=docker&related=docker)

[[IMG]](https://facebook.com/sharer.php?u=https://training.play-with-docker.com/)

[[IMG]](https://plus.google.com/share?url=https://training.play-with-docker.com/)

[[IMG]](http://www.linkedin.com/shareArticle?mini=true&url=https://training.play-with-docker.com/&title=Play%20with%20Docker%20Classroom&source=https://training.play-with-docker.com)

[[IMG]](https://www.docker.com/dockercon/)

[Sign up today](https://www.docker.com/dockercon/)

[Register here](https://dockr.ly/slack)

[here](https://www.docker.com/legal/docker-terms-service)

[[IMG]](https://www.docker.com)

[[IMG]](https://www.facebook.com/docker.run)

[[IMG]](https://twitter.com/docker)

[[IMG]](https://www.github.com/play-with-docker/play-with-docker.github.io)

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container run -it --rm linkextractor:step1 https://training.play-with-docker.com/

/

/about/

#ops

#dev

/ops-stage1

/ops-stage2

/ops-stage3

/dev-stage1

/dev-stage2

/dev-stage3

/alacart

https://twitter.com/intent/tweet?text=Play with Docker Classroom&url=https://training.play-with-docker.com/&via=docker&related=docker

https://facebook.com/sharer.php?u=https://training.play-with-docker.com/

https://plus.google.com/share?url=https://training.play-with-docker.com/

http://www.linkedin.com/shareArticle?mini=true&url=https://training.play-with-docker.com/&title=Play%20with%20Docker%20Classroom&source=https://training.play-with-docker.com

https://www.docker.com/dockercon/

https://www.docker.com/dockercon/

https://dockr.ly/slack

https://www.docker.com/legal/docker-terms-service

https://www.docker.com

https://www.facebook.com/docker.run

https://twitter.com/docker

https://www.github.com/play-with-docker/play-with-docker.github.io

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step3

Branch 'step3' set up to track remote branch 'step3' from 'origin'.

Switched to a new branch 'step3'

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── Dockerfile

├── README.md

├── linkextractor.py

├── main.py

└── requirements.txt

0 directories, 5 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat Dockerfile

FROM python:3

LABEL maintainer="Sawood Alam <@ibnesayeed>"

WORKDIR /app

COPY requirements.txt /app/

RUN pip install -r requirements.txt

COPY \*.py /app/

RUN chmod a+x \*.py

CMD ["./main.py"]

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat main.py

#!/usr/bin/env python

from flask import Flask

from flask import request

from flask import jsonify

from linkextractor import extract\_links

app = Flask(\_\_name\_\_)

@app.route("/")

def index():

return "Usage: http://<hostname>[:<prt>]/api/<url>"

@app.route("/api/<path:url>")

def api(url):

qs = request.query\_string.decode("utf-8")

if qs != "":

url += "?" + qs

links = extract\_links(url)

return jsonify(links)

app.run(host="0.0.0.0")

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker image build -t linkextractor:step3 .

Sending build context to Docker daemon 115.2kB

Step 1/8 : FROM python:3

---> a5d7930b60cc

Step 2/8 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 7d018a9f0c68

Step 3/8 : WORKDIR /app

---> Running in 8a60aedb3494

Removing intermediate container 8a60aedb3494

---> 869ef187a8b0

Step 4/8 : COPY requirements.txt /app/

---> ddfe5d0eb3e4

Step 5/8 : RUN pip install -r requirements.txt

---> Running in 8487f79340d7

docker container run -d -p 5000:5000 --name=linkextractor linkextractor:step3

Collecting beautifulsoup4

Downloading beautifulsoup4-4.10.0-py3-none-any.whl (97 kB)

Collecting flask

Downloading Flask-2.0.2-py3-none-any.whl (95 kB)

Collecting requests

Downloading requests-2.27.1-py2.py3-none-any.whl (63 kB)

Collecting soupsieve>1.2

Downloading soupsieve-2.3.1-py3-none-any.whl (37 kB)

Collecting click>=7.1.2

Downloading click-8.0.3-py3-none-any.whl (97 kB)

Collecting Werkzeug>=2.0

Downloading Werkzeug-2.0.2-py3-none-any.whl (288 kB)

Collecting itsdangerous>=2.0

Downloading itsdangerous-2.0.1-py3-none-any.whl (18 kB)

Collecting Jinja2>=3.0

Downloading Jinja2-3.0.3-py3-none-any.whl (133 kB)

Collecting urllib3<1.27,>=1.21.1

Downloading urllib3-1.26.8-py2.py3-none-any.whl (138 kB)

Collecting certifi>=2017.4.17

Downloading certifi-2021.10.8-py2.py3-none-any.whl (149 kB)

Collecting idna<4,>=2.5

Downloading idna-3.3-py3-none-any.whl (61 kB)

Collecting charset-normalizer~=2.0.0

Downloading charset\_normalizer-2.0.10-py3-none-any.whl (39 kB)

Collecting MarkupSafe>=2.0

Downloading MarkupSafe-2.0.1-cp310-cp310-manylinux\_2\_5\_x86\_64.manylinux1\_x86\_64.manylinux\_2\_12\_x86\_64.manylinux2010\_x86\_64.whl (30 kB)

Installing collected packages: MarkupSafe, Werkzeug, urllib3, soupsieve, Jinja2,itsdangerous, idna, click, charset-normalizer, certifi, requests, flask, beautifulsoup4

Successfully installed Jinja2-3.0.3 MarkupSafe-2.0.1 Werkzeug-2.0.2 beautifulsoup4-4.10.0 certifi-2021.10.8 charset-normalizer-2.0.10 click-8.0.3 flask-2.0.2 idna-3.3 itsdangerous-2.0.1 requests-2.27.1 soupsieve-2.3.1 urllib3-1.26.8

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv

WARNING: You are using pip version 21.2.4; however, version 21.3.1 is available.

You should consider upgrading via the '/usr/local/bin/python -m pip install --upgrade pip' command.

Removing intermediate container 8487f79340d7

---> 9a0b0bf4b752

Step 6/8 : COPY \*.py /app/

---> 9d526ed40b0c

Step 7/8 : RUN chmod a+x \*.py

---> Running in b49279678a1c

Removing intermediate container b49279678a1c

---> 2467e6494a90

Step 8/8 : CMD ["./main.py"]

---> Running in c9eb4ea79495

Removing intermediate container c9eb4ea79495

---> 497b1a9ed6bf

Successfully built 497b1a9ed6bf

Successfully tagged linkextractor:step3

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container run -d -p 5000:5000 --name=linkextractor linkextractor:step3

097f6b5106347504234e7aa8dc531ad7b07da1366a7d0d53217586b6d08b394f

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container ls

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

097f6b510634 linkextractor:step3 "./main.py" 2 seconds ago Up 2 seconds 0.0.0.0:5000->5000/tcp linkextractor

[node1] (local) root@192.168.0.28 ~/linkextractor

$ curl -i http://localhost:5000/api/http://example.com/

HTTP/1.0 200 OK

Content-Type: application/json

Content-Length: 79

Server: Werkzeug/2.0.2 Python/3.10.1

Date: Sat, 08 Jan 2022 04:47:20 GMT

[{"href":"https://www.iana.org/domains/example","text":"More information..."}]

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container logs linkextractor

\* Serving Flask app 'main' (lazy loading)

\* Environment: production

WARNING: This is a development server. Do not use it in a production deployment.

Use a production WSGI server instead.

\* Debug mode: off

\* Running on all addresses.

WARNING: This is a development server. Do not use it in a production deployment.

\* Running on http://172.17.0.2:5000/ (Press CTRL+C to quit)

172.17.0.1 - - [08/Jan/2022 04:47:20] "GET /api/http://example.com/ HTTP/1.1" 200 -

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container rm -f linkextractor

linkextractor

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step4

Branch 'step4' set up to track remote branch 'step4' from 'origin'.

Switched to a new branch 'step4'

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── README.md

├── api

│ ├── Dockerfile

│ ├── linkextractor.py

│ ├── main.py

│ └── requirements.txt

├── docker-compose.yml

└── www

└── index.php

2 directories, 7 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat docker-compose.yml

version: '3'

services:

api:

image: linkextractor-api:step4-python

build: ./api

ports:

- "5000:5000"

web:

image: php:7-apache

ports:

- "80:80"

environment:

- API\_ENDPOINT=http://api:5000/api/

volumes:

- ./www:/var/www/html

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat www/index.php

<!DOCTYPE html>

<?php

$api\_endpoint = $\_ENV["API\_ENDPOINT"] ?: "http://localhost:5000/api/";

$url = "";

if(isset($\_GET["url"]) && $\_GET["url"] != "") {

$url = $\_GET["url"];

$json = @file\_get\_contents($api\_endpoint . $url);

if($json == false) {

$err = "Something is wrong with the URL: " . $url;

} else {

$links = json\_decode($json, true);

$domains = [];

foreach($links as $link) {

array\_push($domains, parse\_url($link["href"], PHP\_URL\_HOST));

}

$domainct = @array\_count\_values($domains);

arsort($domainct);

}

}

?>

<html>

<head>

<meta charset="utf-8">

<title>Link Extractor</title>

<style media="screen">

html {

background: #EAE7D6;

font-family: sans-serif;

}

body {

margin: 0;

}

h1 {

padding: 10px;

margin: 0 auto;

color: #EAE7D6;

max-width: 600px;

}

h1 a {

text-decoration: none;

color: #EAE7D6;

}

h2 {

background: #082E41;

color: #EAE7D6;

margin: -10px;

padding: 10px;

}

p {

margin: 25px 5px 5px;

}

section {

max-width: 600px;

margin: 10px auto;

padding: 10px;

border: 1px solid #082E41;

}

div.header {

background: #082E41;

margin: 0;

}

div.footer {

background: #082E41;

margin: 0;

padding: 5px;

}

.footer p {

margin: 0 auto;

max-width: 600px;

color: #EAE7D6;

text-align: center;

}

.footer p a {

color: #24C2CB;

text-decoration: none;

}

.error {

color: #DA2536;

}

form {

display: flex;

}

input {

font-size: 20px;

padding: 3px;

height: 40px;

}

input.text {

box-sizing:border-box;

flex-grow: 1;

border-color: #082E41;

}

input.button {

width: 150px;

background: #082E41;

border-color: #082E41;

color: #EAE7D6;

}

table {

width: 100%;

text-align: left;

margin-top: 10px;

}

table th, table td {

padding: 3px;

}

table th:last-child, table td:last-child {

width: 70px;

text-align: right;

}

table th {

border-top: 1px solid #082E41;

border-bottom: 1px solid #082E41;

}

table tr:last-child td {

border-top: 1px solid #082E41;

border-bottom: 1px solid #082E41;

}

</style>

</head>

<body>

<div class="header">

<h1><a href="/">Link Extractor</a></h1>

</div>

<section>

<form action="/">

<input class="text" type="text" name="url" placeholder="http://example.com/" value="<?php echo $url; ?>">

<input class="button" type="submit" value="Extract Links">

</form>

</section>

<?php if(isset($err)): ?>

<section>

<h2>Error</h2>

<p class="error"><?php echo $err; ?></p>

</section>

<?php endif; ?>

<?php if($url != "" && !isset($err)): ?>

<section>

<h2>Summary</h2>

<p>

<strong>Page:</strong> <?php echo "<a href=\"" . $url . "\">" . $url ."</a>"; ?>

</p>

<table>

<tr>

<th>Domain</th>

<th># Links</th>

</tr>

<?php

foreach($domainct as $key => $value) {

echo "<tr>";

echo "<td>" . $key . "</td>";

echo "<td>" . $value . "</td>";

echo "</tr>";

}

?>

<tr>

<td><strong>Total</strong></td>

<td><strong><?php echo count($links); ?></strong></td>

</tr>

</table>

</section>

<section>

<h2>Links</h2>

<ul>

<?php

foreach($links as $link) {

echo "<li><a href=\"" . $link["href"] . "\">" . $link["text"] . "</a></li>";

}

?>

</ul>

</section>

<?php endif; ?>

<div class="footer">

<p><a href="https://github.com/ibnesayeed/linkextractor">Link Extractor</a> by <a href="https://twitter.com/ibnesayeed">@ibnesayeed</a> from

<a href="https://ws-dl.cs.odu.edu/">WS-DL, ODU</a>

</p>

</div>

</body>

</html>

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose up -d --build

Creating network "linkextractor\_default" with the default driver

Building api

Step 1/8 : FROM python:3

---> a5d7930b60cc

Step 2/8 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 7d018a9f0c68

Step 3/8 : WORKDIR /app

---> Using cache

---> 869ef187a8b0

Step 4/8 : COPY requirements.txt /app/

---> 3093c42fac7e

Step 5/8 : RUN pip install -r requirements.txt

---> Running in a7de6417bb5d

Collecting beautifulsoup4

Downloading beautifulsoup4-4.10.0-py3-none-any.whl (97 kB)

Collecting flask

Downloading Flask-2.0.2-py3-none-any.whl (95 kB)

Collecting requests

Downloading requests-2.27.1-py2.py3-none-any.whl (63 kB)

Collecting soupsieve>1.2

Downloading soupsieve-2.3.1-py3-none-any.whl (37 kB)

Collecting Jinja2>=3.0

Downloading Jinja2-3.0.3-py3-none-any.whl (133 kB)

Collecting itsdangerous>=2.0

Downloading itsdangerous-2.0.1-py3-none-any.whl (18 kB)

Collecting Werkzeug>=2.0

Downloading Werkzeug-2.0.2-py3-none-any.whl (288 kB)

Collecting click>=7.1.2

Downloading click-8.0.3-py3-none-any.whl (97 kB)

Collecting charset-normalizer~=2.0.0

Downloading charset\_normalizer-2.0.10-py3-none-any.whl (39 kB)

Collecting idna<4,>=2.5

Downloading idna-3.3-py3-none-any.whl (61 kB)

Collecting urllib3<1.27,>=1.21.1

Downloading urllib3-1.26.8-py2.py3-none-any.whl (138 kB)

Collecting certifi>=2017.4.17

Downloading certifi-2021.10.8-py2.py3-none-any.whl (149 kB)

Collecting MarkupSafe>=2.0

Downloading MarkupSafe-2.0.1-cp310-cp310-manylinux\_2\_5\_x86\_64.manylinux1\_x86\_64.manylinux\_2\_12\_x86\_64.manylinux2010\_x86\_64.whl (30 kB)

Installing collected packages: MarkupSafe, Werkzeug, urllib3, soupsieve, Jinja2,itsdangerous, idna, click, charset-normalizer, certifi, requests, flask, beautifulsoup4

Successfully installed Jinja2-3.0.3 MarkupSafe-2.0.1 Werkzeug-2.0.2 beautifulsoup4-4.10.0 certifi-2021.10.8 charset-normalizer-2.0.10 click-8.0.3 flask-2.0.2 idna-3.3 itsdangerous-2.0.1 requests-2.27.1 soupsieve-2.3.1 urllib3-1.26.8

WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv

WARNING: You are using pip version 21.2.4; however, version 21.3.1 is available.

You should consider upgrading via the '/usr/local/bin/python -m pip install --upgrade pip' command.

Removing intermediate container a7de6417bb5d

---> 11ce438180fe

Step 6/8 : COPY \*.py /app/

---> b154648b3a1a

Step 7/8 : RUN chmod a+x \*.py

---> Running in b71647bf3853

Removing intermediate container b71647bf3853

---> 743a942297c9

Step 8/8 : CMD ["./main.py"]

---> Running in 5de137296e98

Removing intermediate container 5de137296e98

---> 1be21fd11744

Successfully built 1be21fd11744

Successfully tagged linkextractor-api:step4-python

Pulling web (php:7-apache)...

7-apache: Pulling from library/php

a2abf6c4d29d: Pull complete

c5608244554d: Pull complete

2d07066487a0: Pull complete

1b6dfaf1958c: Pull complete

32c5e6a60073: Pull complete

90cf855b27cc: Pull complete

8b0f1068c586: Pull complete

5355461305e8: Pull complete

ad1eec592342: Pull complete

e03fbc76cb78: Pull complete

1f5796e48b39: Pull complete

72fbe8e1d4e7: Pull complete

96edece66175: Pull complete

Digest: sha256:729ad01c7d8e10fd992a6d4f3eb05dce3fb69bdf5c4fb4a9de4be4f4f5ae4dcc

Status: Downloaded newer image for php:7-apache

Creating linkextractor\_api\_1 ... done

Creating linkextractor\_web\_1 ... done

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container ls

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

2fd95fab5d64 php:7-apache "docker-php-entrypoi…" 2 seconds ago Up 1 second 0.0.0.0:80->80/tcp linkextractor\_web\_1

c865fc4b2115 linkextractor-api:step4-python "./main.py" 2 seconds ago Up 1 second 0.0.0.0:5000->5000/tcp linkextractor\_api\_1

[node1] (local) root@192.168.0.28 ~/linkextractor

$ curl -i http://localhost:5000/api/http://example.com/

HTTP/1.0 200 OK

Content-Type: application/json

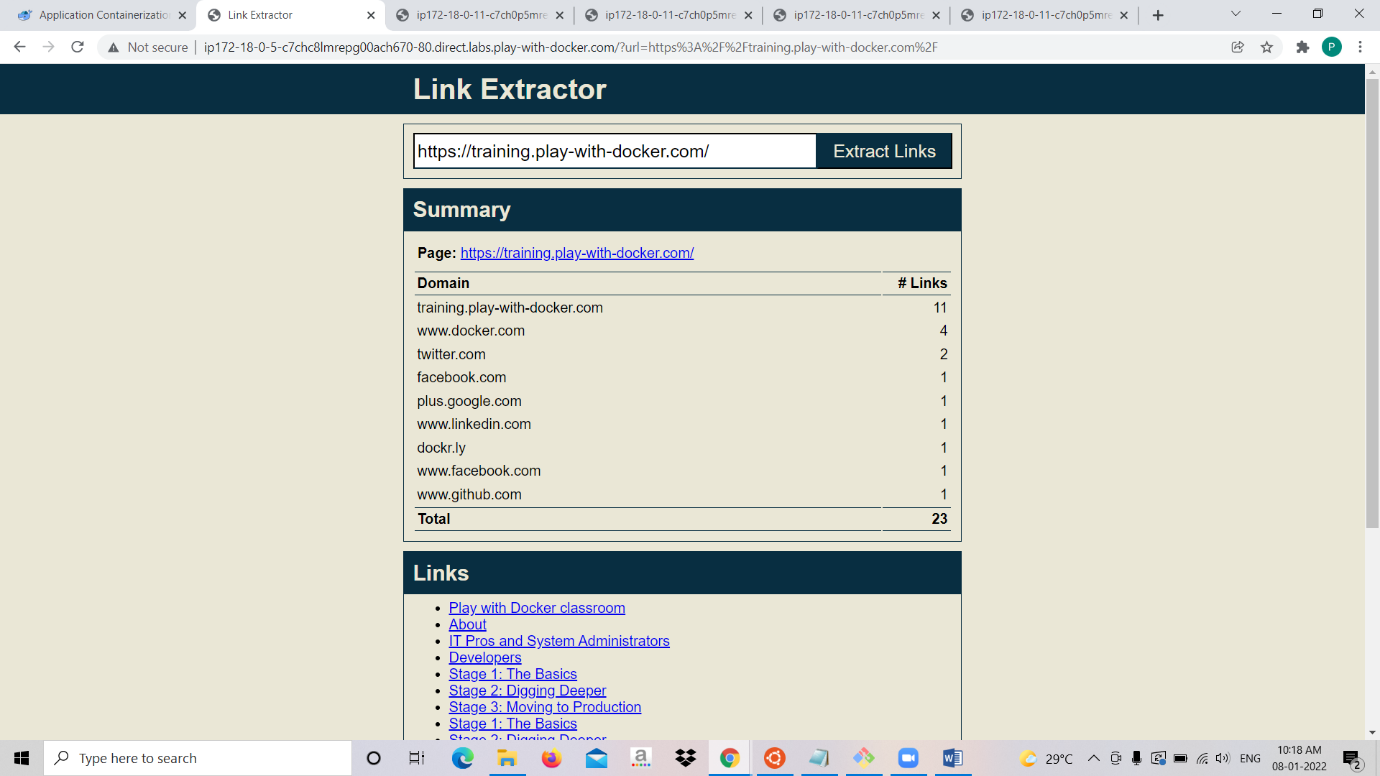
Content-Length: 79

Server: Werkzeug/2.0.2 Python/3.10.1

Date: Sat, 08 Jan 2022 04:49:30 GMT

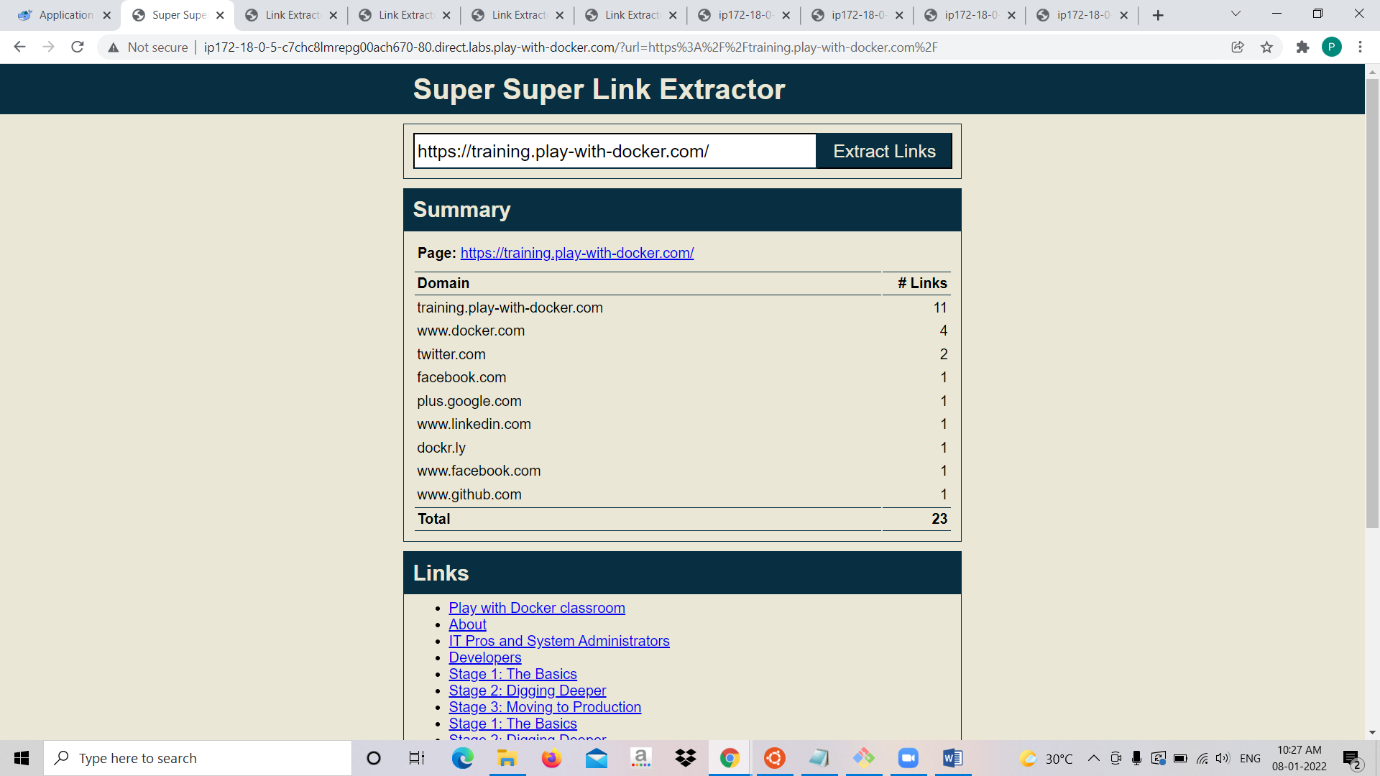
[{"href":"https://www.iana.org/domains/example","text":"More information..."}]

[node1] (local) root@192.168.0.28 ~/linkextractor



[node1] (local) root@192.168.0.28 ~/linkextractor

$ sed -i 's/Link Extractor/Super Link Extractor/g' www/index.php



[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step4

M www/index.php

Switched to branch 'step4'

Your branch is up to date with 'origin/step4'.

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── README.md

├── api

│ ├── Dockerfile

│ ├── linkextractor.py

│ ├── main.py

│ └── requirements.txt

├── docker-compose.yml

├── logs

│ └── extraction.log

└── www

└── index.php

3 directories, 8 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat docker-compose.yml

version: '3'

services:

api:

image: linkextractor-api:step4-python

build: ./api

ports:

- "5000:5000"

web:

image: php:7-apache

ports:

- "80:80"

environment:

- API\_ENDPOINT=http://api:5000/api/

volumes:

- ./www:/var/www/html

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat www/index.php

<!DOCTYPE html>

<?php

$api\_endpoint = $\_ENV["API\_ENDPOINT"] ?: "http://localhost:5000/api/";

$url = "";

if(isset($\_GET["url"]) && $\_GET["url"] != "") {

$url = $\_GET["url"];

$json = @file\_get\_contents($api\_endpoint . $url);

if($json == false) {

$err = "Something is wrong with the URL: " . $url;

} else {

$links = json\_decode($json, true);

$domains = [];

foreach($links as $link) {

array\_push($domains, parse\_url($link["href"], PHP\_URL\_HOST));

}

$domainct = @array\_count\_values($domains);

arsort($domainct);

}

}

?>

<html>

<head>

<meta charset="utf-8">

<title>Super Link Extractor</title>

<style media="screen">

html {

background: #EAE7D6;

font-family: sans-serif;

}

body {

margin: 0;

}

h1 {

padding: 10px;

margin: 0 auto;

color: #EAE7D6;

max-width: 600px;

}

h1 a {

text-decoration: none;

color: #EAE7D6;

}

h2 {

background: #082E41;

color: #EAE7D6;

margin: -10px;

padding: 10px;

}

p {

margin: 25px 5px 5px;

}

section {

max-width: 600px;

margin: 10px auto;

padding: 10px;

border: 1px solid #082E41;

}

div.header {

background: #082E41;

margin: 0;

}

div.footer {

background: #082E41;

margin: 0;

padding: 5px;

}

.footer p {

margin: 0 auto;

max-width: 600px;

color: #EAE7D6;

text-align: center;

}

.footer p a {

color: #24C2CB;

text-decoration: none;

}

.error {

color: #DA2536;

}

form {

display: flex;

}

input {

font-size: 20px;

padding: 3px;

height: 40px;

}

input.text {

box-sizing:border-box;

flex-grow: 1;

border-color: #082E41;

}

input.button {

width: 150px;

background: #082E41;

border-color: #082E41;

color: #EAE7D6;

}

table {

width: 100%;

text-align: left;

margin-top: 10px;

}

table th, table td {

padding: 3px;

}

table th:last-child, table td:last-child {

width: 70px;

text-align: right;

}

table th {

border-top: 1px solid #082E41;

border-bottom: 1px solid #082E41;

}

table tr:last-child td {

border-top: 1px solid #082E41;

border-bottom: 1px solid #082E41;

}

</style>

</head>

<body>

<div class="header">

<h1><a href="/">Super Link Extractor</a></h1>

</div>

<section>

<form action="/">

<input class="text" type="text" name="url" placeholder="http://example.com/" value="<?php echo $url; ?>">

<input class="button" type="submit" value="Extract Links">

</form>

</section>

<?php if(isset($err)): ?>

<section>

<h2>Error</h2>

<p class="error"><?php echo $err; ?></p>

</section>

<?php endif; ?>

<?php if($url != "" && !isset($err)): ?>

<section>

<h2>Summary</h2>

<p>

<strong>Page:</strong> <?php echo "<a href=\"" . $url . "\">" . $url ."</a>"; ?>

</p>

<table>

<tr>

<th>Domain</th>

<th># Links</th>

</tr>

<?php

foreach($domainct as $key => $value) {

echo "<tr>";

echo "<td>" . $key . "</td>";

echo "<td>" . $value . "</td>";

echo "</tr>";

}

?>

<tr>

<td><strong>Total</strong></td>

<td><strong><?php echo count($links); ?></strong></td>

</tr>

</table>

</section>

<section>

<h2>Links</h2>

<ul>

<?php

foreach($links as $link) {

echo "<li><a href=\"" . $link["href"] . "\">" . $link["text"] . "</a></li>";

}

?>

</ul>

</section>

<?php endif; ?>

<div class="footer">

<p><a href="https://github.com/ibnesayeed/linkextractor">Super Link Extractor</a> by <a href="https://twitter.com/ibnesayeed">@ibnesayeed</a> from

<a href="https://ws-dl.cs.odu.edu/">WS-DL, ODU</a>

</p>

</div>

</body>

</html>

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose up -d --build

WARNING: Found orphan containers (linkextractor\_redis\_1) for this project. If you removed or renamed this service in your compose file, you can run this command with the --remove-orphans flag to clean it up.

Building api

Step 1/8 : FROM python:3

---> a5d7930b60cc

Step 2/8 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 7d018a9f0c68

Step 3/8 : WORKDIR /app

---> Using cache

---> 869ef187a8b0

Step 4/8 : COPY requirements.txt /app/

---> Using cache

---> 3093c42fac7e

Step 5/8 : RUN pip install -r requirements.txt

---> Using cache

---> 11ce438180fe

Step 6/8 : COPY \*.py /app/

---> Using cache

---> b154648b3a1a

Step 7/8 : RUN chmod a+x \*.py

---> Using cache

---> 743a942297c9

Step 8/8 : CMD ["./main.py"]

---> Using cache

---> 1be21fd11744

Successfully built 1be21fd11744

Successfully tagged linkextractor-api:step4-python

Recreating linkextractor\_api\_1 ... done

Recreating linkextractor\_web\_1 ... done

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker container ls

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

dd3935ead8f9 php:7-apache "docker-php-entrypoi…" 3 seconds ago Up 1 second 0.0.0.0:80->80/tcp linkextractor\_web\_1

52cd478ab402 linkextractor-api:step4-python "./main.py" 4 seconds ago Up 2 seconds 0.0.0.0:5000->5000/tcp linkextractor\_api\_1

602231cf1a4a redis "docker-entrypoint.s…" 5 minutes ago Up 5 minutes 6379/tcp linkextractor\_redis\_1

[node1] (local) root@192.168.0.28 ~/linkextractor

$ curl -i http://localhost:5000/api/http://example.com/

HTTP/1.0 200 OK

Content-Type: application/json

Content-Length: 79

Server: Werkzeug/2.0.2 Python/3.10.1

Date: Sat, 08 Jan 2022 04:58:06 GMT

[{"href":"https://www.iana.org/domains/example","text":"More information..."}]

[node1] (local) root@192.168.0.28 ~/linkextractor

$ sed -i 's/Link Extractor/Super Link Extractor/g' www/index.php

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git reset --hard

HEAD is now at 2a3ec3e Synchronize branch step4

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose down

Stopping linkextractor\_web\_1 ... done

Stopping linkextractor\_api\_1 ...

git checkout step5

tree

cat www/Dockerfile

Stopping linkextractor\_api\_1 ... done

WARNING: Found orphan containers (linkextractor\_redis\_1) for this project. If you removed or renamed this service in your compose file, you can run this command with the --remove-orphans flag to clean it up.

Removing linkextractor\_web\_1 ... done

Removing linkextractor\_api\_1 ... done

Removing network linkextractor\_default

ERROR: error while removing network: network linkextractor\_default id 6da937af00f142c707af634004e691a6942d6a1bc3bedca144befa024b8439f9 has active endpoints

[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step5

Switched to branch 'step5'

Your branch is up to date with 'origin/step5'.

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── README.md

├── api

│ ├── Dockerfile

│ ├── linkextractor.py

│ ├── main.py

│ └── requirements.txt

├── docker-compose.yml

├── logs

│ └── extraction.log

└── www

├── Dockerfile

└── index.php

3 directories, 9 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat www/Dockerfile

FROM php:7-apache

LABEL maintainer="Sawood Alam <@ibnesayeed>"

ENV API\_ENDPOINT="http://localhost:5000/api/"

COPY . /var/www/html/

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat api/main.py

#!/usr/bin/env python

import os

import json

import redis

from flask import Flask

from flask import request

from linkextractor import extract\_links

app = Flask(\_\_name\_\_)

redis\_conn = redis.from\_url(os.getenv("REDIS\_URL", "redis://localhost:6379"))

@app.route("/")

def index():

return "Usage: http://<hostname>[:<prt>]/api/<url>"

@app.route("/api/<path:url>")

def api(url):

qs = request.query\_string.decode("utf-8")

if qs != "":

url += "?" + qs

jsonlinks = redis\_conn.get(url)

if not jsonlinks:

links = extract\_links(url)

jsonlinks = json.dumps(links, indent=2)

redis\_conn.set(url, jsonlinks)

response = app.response\_class(

status=200,

mimetype="application/json",

response=jsonlinks

)

return response

app.run(host="0.0.0.0")

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat docker-compose.yml

version: '3'

services:

api:

image: linkextractor-api:step5-python

build: ./api

ports:

- "5000:5000"

environment:

- REDIS\_URL=redis://redis:6379

web:

image: linkextractor-web:step5-php

build: ./www

ports:

- "80:80"

environment:

- API\_ENDPOINT=http://api:5000/api/

redis:

image: redis

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose up -d --build

Building api

Step 1/9 : FROM python:3

---> a5d7930b60cc

Step 2/9 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 7d018a9f0c68

Step 3/9 : ENV REDIS\_URL="redis://localhost:6379"

---> Using cache

---> 6694938c42b9

Step 4/9 : WORKDIR /app

---> Using cache

---> 284fce056391

Step 5/9 : COPY requirements.txt /app/

---> Using cache

---> 128666908e8c

Step 6/9 : RUN pip install -r requirements.txt

---> Using cache

---> 30a6bdc3542a

Step 7/9 : COPY \*.py /app/

---> Using cache

---> e746d8d67703

Step 8/9 : RUN chmod a+x \*.py

---> Using cache

---> 7db3d71dfede

Step 9/9 : CMD ["./main.py"]

---> Using cache

---> 9a1e8b153aeb

Successfully built 9a1e8b153aeb

Successfully tagged linkextractor-api:step5-python

Building web

Step 1/4 : FROM php:7-apache

---> 0a4f19d60710

Step 2/4 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> e27666b3dd34

Step 3/4 : ENV API\_ENDPOINT="http://localhost:5000/api/"

---> Using cache

---> 5e4b8617dd07

Step 4/4 : COPY . /var/www/html/

---> Using cache

---> 6b8c9577cb18

Successfully built 6b8c9577cb18

Successfully tagged linkextractor-web:step5-php

linkextractor\_redis\_1 is up-to-date

Creating linkextractor\_web\_1 ... done

Creating linkextractor\_api\_1 ... done

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose exec redis redis-cli monitor

OK

sed -i 's/Link Extractor/Super Link Extractor/g' www/index.php

git reset --hard

docker-compose down

git checkout step6

tree

^C[node1] (local) root@192.168.0.28 ~/linkextractor

$ git checkout step6

Switched to branch 'step6'

Your branch is up to date with 'origin/step6'.

[node1] (local) root@192.168.0.28 ~/linkextractor

$ tree

.

├── README.md

├── api

│ ├── Dockerfile

│ ├── Gemfile

│ └── linkextractor.rb

├── docker-compose.yml

├── logs

│ └── extraction.log

└── www

├── Dockerfile

└── index.php

3 directories, 8 files

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat api/linkextractor.rb

#!/usr/bin/env ruby

# encoding: utf-8

require "sinatra"

require "open-uri"

require "uri"

require "nokogiri"

require "json"

require "redis"

set :protection, :except=>:path\_traversal

redis = Redis.new(url: ENV["REDIS\_URL"] || "redis://localhost:6379")

Dir.mkdir("logs") unless Dir.exist?("logs")

cache\_log = File.new("logs/extraction.log", "a")

get "/" do

"Usage: http://<hostname>[:<prt>]/api/<url>"

end

get "/api/\*" do

url = [params['splat'].first, request.query\_string].reject(&:empty?).join("?")

cache\_status = "HIT"

jsonlinks = redis.get(url)

if jsonlinks.nil?

cache\_status = "MISS"

jsonlinks = JSON.pretty\_generate(extract\_links(url))

redis.set(url, jsonlinks)

end

cache\_log.puts "#{Time.now.to\_i}\t#{cache\_status}\t#{url}"

status 200

headers "content-type" => "application/json"

body jsonlinks

end

def extract\_links(url)

links = []

doc = Nokogiri::HTML(open(url))

doc.css("a").each do |link|

text = link.text.strip.split.join(" ")

begin

links.push({

text: text.empty? ? "[IMG]" : text,

href: URI.join(url, link["href"])

})

rescue

end

end

links

end

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat api/Dockerfile

FROM ruby:2.6

LABEL maintainer="Sawood Alam <@ibnesayeed>"

ENV LANG C.UTF-8

ENV REDIS\_URL="redis://localhost:6379"

WORKDIR /app

COPY Gemfile /app/

RUN bundle install

COPY linkextractor.rb /app/

RUN chmod a+x linkextractor.rb

CMD ["./linkextractor.rb", "-o", "0.0.0.0"]

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat docker-compose.yml

version: '3'

services:

api:

image: linkextractor-api:step6-ruby

build: ./api

ports:

- "4567:4567"

environment:

- REDIS\_URL=redis://redis:6379

volumes:

- ./logs:/app/logs

web:

image: linkextractor-web:step6-php

build: ./www

ports:

- "80:80"

environment:

- API\_ENDPOINT=http://api:4567/api/

redis:

image: redis

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose up -d --build

Building api

Step 1/10 : FROM ruby:2.6

---> 9a2d45e2219e

Step 2/10 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> 4e7bfd34ee82

Step 3/10 : ENV LANG C.UTF-8

---> Using cache

---> fcf4df144d17

Step 4/10 : ENV REDIS\_URL="redis://localhost:6379"

---> Using cache

---> 2e73946414b9

Step 5/10 : WORKDIR /app

---> Using cache

---> daba9d2f1792

Step 6/10 : COPY Gemfile /app/

---> Using cache

---> 2d070639befb

Step 7/10 : RUN bundle install

---> Using cache

---> ae37543108a0

Step 8/10 : COPY linkextractor.rb /app/

---> Using cache

---> 2dad891681fd

Step 9/10 : RUN chmod a+x linkextractor.rb

---> Using cache

---> d944a5911498

Step 10/10 : CMD ["./linkextractor.rb", "-o", "0.0.0.0"]

---> Using cache

---> b99017e89e6e

Successfully built b99017e89e6e

Successfully tagged linkextractor-api:step6-ruby

Building web

Step 1/3 : FROM php:7-apache

---> 0a4f19d60710

Step 2/3 : LABEL maintainer="Sawood Alam <@ibnesayeed>"

---> Using cache

---> e27666b3dd34

Step 3/3 : COPY . /var/www/html/

---> Using cache

---> 691d0fca50aa

Successfully built 691d0fca50aa

Successfully tagged linkextractor-web:step6-php

Recreating linkextractor\_web\_1 ...

Recreating linkextractor\_web\_1 ... done

Recreating linkextractor\_api\_1 ...

Recreating linkextractor\_api\_1 ... done

[node1] (local) root@192.168.0.28 ~/linkextractor

$ curl -i http://localhost:4567/api/http://example.com/

curl: (56) Recv failure: Connection reset by peer

[node1] (local) root@192.168.0.28 ~/linkextractor

$ docker-compose down

Stopping linkextractor\_api\_1 ... done

Stopping linkextractor\_web\_1 ... done

Stopping linkextractor\_redis\_1 ... done

Removing linkextractor\_api\_1 ... done

Removing linkextractor\_web\_1 ... done

Removing linkextractor\_redis\_1 ... done

Removing network linkextractor\_default

[node1] (local) root@192.168.0.28 ~/linkextractor

$ cat logs/extraction.log

1641617620 MISS http://example.com/