Docker Meetup Tokyo #35 (2021/5/28)

Docker 20.10

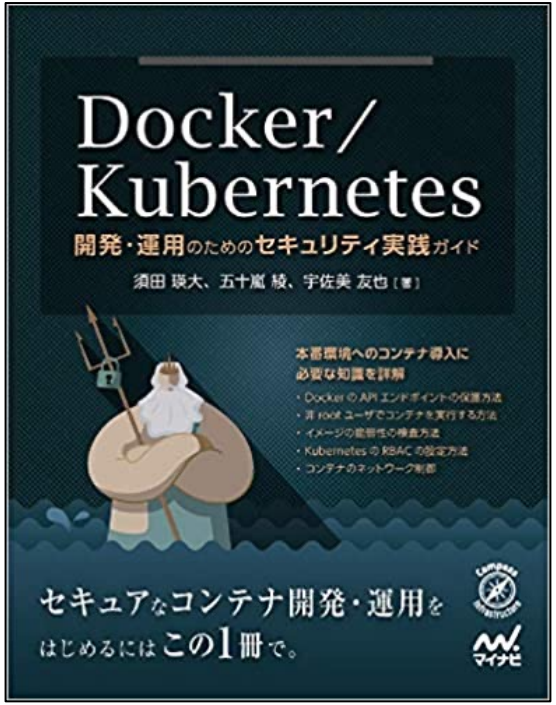
日本電信電話株式会社
ソフトウェアイノベーションセンタ
須田 瑛大

自己紹介

- コンテナ関連OSSのメンテナ (コミッタ)

  Moby (OSS版Docker), BuildKit, containerd, runc, Rootlessコンテナ関連など

- 書籍「Docker/Kubernetes 開発・運用のためのセキュリティ実践ガイド」執筆
  https://www.amazon.co.jp/dp/4839970505



- その他 色々 https://github.com/AkihiroSuda

Docker 20.10

- 2020年12月9日にリリース
- Docker 19.03 (2019年7月) 以来、約1年半ぶりのリリース

- RHEL8系のデフォルトの設定で動くようになった
- Fedoraのデフォルトの設定で動くようになった
- RootlessやBuildXなどがexperimentalから正式機能に昇格した
- Swarm Jobsに対応した

RHEL8系のデフォルトで動くようになった

NTT

- RHEL 8系OS (CentOS 8などを含む) のデフォルトでは、Docker 19.03はうまく動かなかった
  - コンテナ内からホスト名を解決できない docker/for-linux#957
  - ポートを公開できない libnetwork#2496
- Dockerがfirewalldに対応していなかったのが原因
- Docker 20.10ではRHEL 8系のデフォルトで動くようになった

RHEL8系のデフォルトで動くようになった

- ただし、RHEL 8のRed Hat社 公式パッケージからは、Dockerは削除されたまま

- Docker社からも、RHEL 8向けのパッケージは提供されていない

- CentOS 8向けには、Docker社 公式パッケージが提供されている
  https://download.docker.com/linux/centos/8/x86_64/stable/Packages/
  - CentOS 8とバイナリ互換の他のOSでも動くはず
    (Alma Linux, Oracle Linux, Rocky Linux, VzLinux...)

Fedoraのデフォルトで動くようになった

NTT

- Fedora 31以降のデフォルトでは、Docker 19.03は動かなかった
- Dockerがcgroup v2に対応していなかったのが原因
  - CPUやメモリなどのリソース制御に使われている、Linuxカーネルの機能の新バージョン
- Docker 20.10ではFedoraのデフォルトで動くようになった
  - 実装としては2019年から動いていたが、リリースにかなり時間がかかった
  - 20.10.0ではexperimental扱い、20.10.6から正式扱い

Rootlessモードがexperimentalから正式に昇格

- ホストのroot権限を使わずにdockerdやコンテナを動かすことで、セキュリティを強化する技術
  https://rootlesscontaine.rs/getting-started/docker/

- 2018年: POC実装
  2019年: Docker 19.03 にて experimental としてマージ
  2020年: Docker 20.10 にて 正式機能に昇格

Rootlessモードがexperimentalから正式に昇格

- CPUやメモリのリソースの制御 (docker run –-cpus, ---memory) に対応した
  - ホストが前述のcgroup v2に対応している必要がある

- aptやdnfでインストールできるようになった
  - sudo apt-get install docker-ce-rootless-extras
    - sudoが要るのは Dockerの制約ではなく apt-getやdnfの制約
  - sudo無し、パッケージマネージャ無しでのインストールも従来通り可能
    - https://get.docker.com/rootless

Swarm Jobs

- 最近名前を聞かないSwarmであるが、開発は続いている

- バッチジョブ型の”Job” serviceも動かせるようになった
  - KubernetesのJobマニフェストに類似する
  - 従来型の service は、KubernetesのDeploymentマニフェストやDaemonSetに類似する

Swarm Jobs

- クラスタ内の全ノードのホストファイルシステムに、/etc/some-file を作成する例

```
$ docker service create ¥
  --name create-file-on-all-nodes ¥
  --mode global-job ¥
  --mount type=bind,src=/,dst=/mnt ¥
  alpine ¥
  sh -exc "echo foo > /mnt/etc/some-file"
```

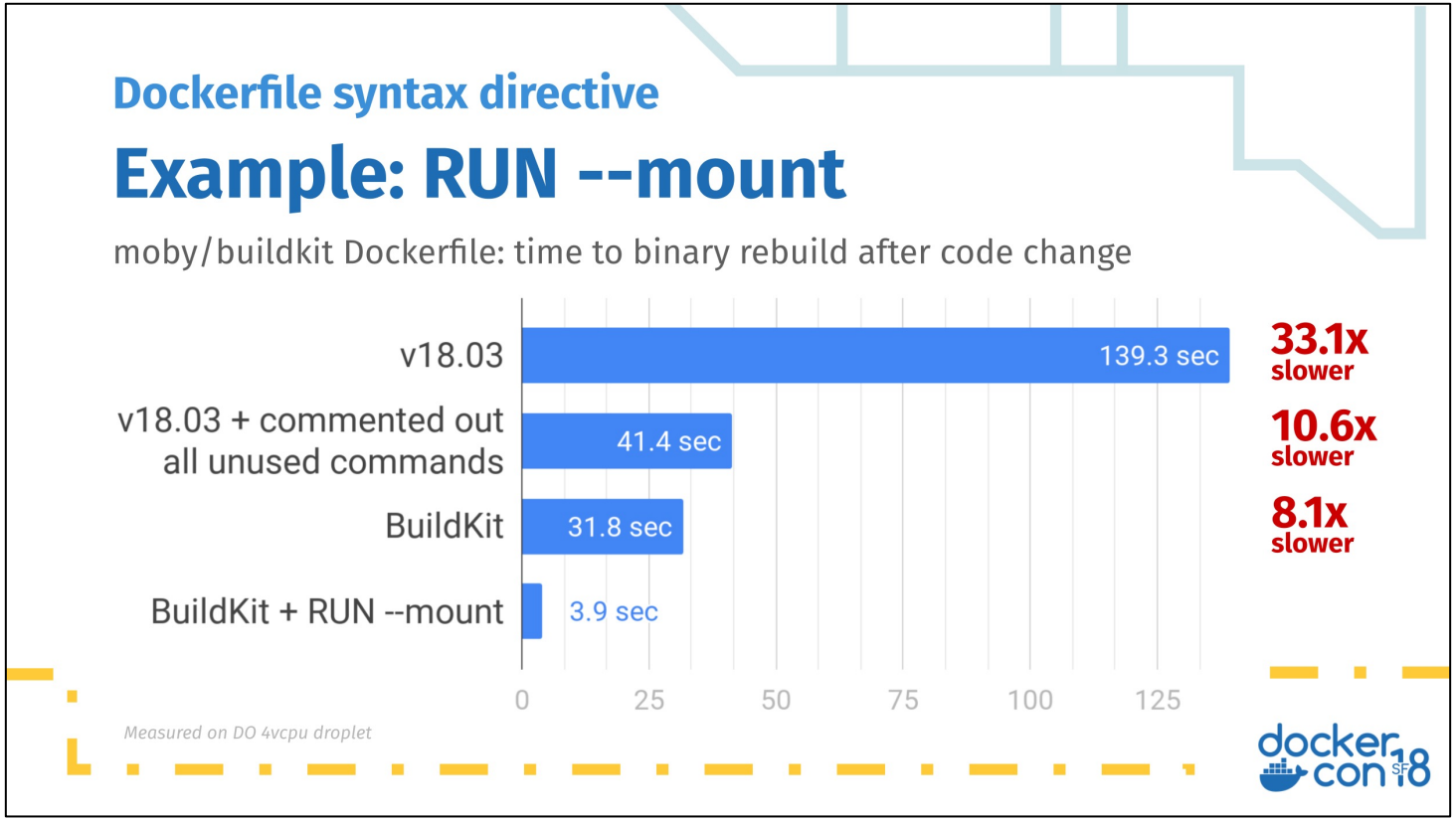
Dockerfile: RUN --mount=type=...が正式機能に昇格

Docker 18.06 にて experimentalとして導入された機能

- RUN –-mount=type=cache
  - apt、maven、npm などのパッケージマネージャや、コンパイラのキャッシュを保持できる機能
- RUN –-mount=type=secret　　(Docker 18.09から)
  - SSHやS3などの認証情報を安全に扱える機能
- RUN –-mount=type=ssh　　(Docker 18.09から)
  - 上記secretのSSH特化版 (パスフレーズ対応など)

など

Dockerfile: RUN --mount=type=...が正式機能に昇格



"Docker Platform Internals: containerd and BuildKit" (DockerCon 2018)
Michael Crosby and Tonis Tiigi, Docker https://t.co/aUKqQCVmXa

Dockerfile: RUN --mount=type=...が正式機能に昇格

- Docker 19.03まででは、BuildKitモードを有効化し、さらにDockerfileの先頭に次の行を書く必要があった

```
# syntax = docker/dockerfile:experimental
```

- Docker 20.10からは # syntax = … 行が不要になった

- BuildKitモードは依然として有効化しておく必要がある
  - Docker for Mac/Winでは 2020年9月からデフォルトで有効

docker buildxが正式機能に昇格

- BuildKitモード (DOCKER_BUILDKIT=1) を更に拡張したモード
  https://github.com/docker/buildx

- マルチプラットフォームイメージのビルドや、Kubernetesを使った分散ビルドなど、docker build ではできない先進的な機能が備わっている

- docker-ce-cli パッケージに正式機能として含まれるようになった

docker buildxが正式機能に昇格

- リモートのARMマシンを使ってマルチプラットフォームイメージをビルドする例

```
$ docker buildx create --name remote --use
$ docker buildx create --name remote ¥
  --append ssh://me@my-arm-instance
$ docker buildx build ¥
  --push -t example.com/hello:latest ¥
  --platform=linux/amd64,linux/arm64 .
```

  - リモートのARMマシンを使わず、QEMUでエミュレートするモードもある

docker buildxが正式機能に昇格

- Kubernetesクラスタを使って分散ビルドする例

```
$ docker buildx create    ¥
  --driver kubernetes     ¥
  --driver-opt replicas=3 ¥
  --use

$ docker buildx build ¥
  --push -t example.com/hello:latest
```

  - 複数のDockerfileを同時にビルドすると負荷が分散される

Dockerfile: COPY –-chmod、ADD --chmod

- COPY --chown、ADD --chown は Docker 17.09から存在
  - コンテナを非rootユーザで動かすときに便利
- Docker 20.10にて、COPY --chmod、ADD --chmod にも対応
- 以前のバージョンでも、COPYしてからRUN chmodすることはできたが、イメージの容量が膨らむ問題があった

docker run --pull=(always|missing|never)

- KubernetesのimagePullPolicyに似た機能

- 確実に最新のイメージを動かしたいとき、pullしてからrunする手間を省ける (docker run --pull=always)

Docker 21.XX に入りそうなpull requests

NTT

- Swarm cluster volumes moby#42404
  - Kubernetesと同じくCSI (Container Storage Interface) を使う
- libnetworkのリポジトリがMoby (dockerd) のリポジトリに統合される moby#42262
  - コントリビュートしやすくなるはず
- Dockerfileでheredocが使えるようになる buildkit#2132
  - <<EOF みたいなやつのこと
  - 複数行にまたがる命令を書きやすくなる

```
RUN <<EOF
  apt-get update
  apt-get install foo
  foo ...
EOF
```

Docker 20.10 まとめ

- RHEL8系のデフォルトの設定で動くようになった
- Fedoraのデフォルトの設定で動くようになった
- RootlessやBuildXなどがexperimentalから正式機能に昇格した
- Swarm Jobsに対応した

- その他、細かい情報はnttlabsブログにまとめました
  https://medium.com/nttlabs/docker-20-10-59cc4bd59d37