GCP のサーバーレスを支える gVisor

酔いどれGCPUG 2018/06/25
@apstndb

誰？

← これ

- 主に GCPUG Slack と Twitter に生息
- GCPUG の #blog-google-cloud-jp 有志と「TWiGCP(This Week in GCP)」の翻訳＋日本語注釈の「今週の GCP」がなんとか続いている

アジェンダ

- appengine ja night 文脈で話した Java 8 ランタイム以降のサンドボックスと gVisor に入れられなかったネタを早口で喋ります

- 話すこと
  - パブリックな情報から推測可能な Google での gVisor 利用
- 話さないこと
  - サーバーレスとは何か
  - App Engine での詳細(もう話した)
  - gVisor の詳細
  - EAP/Closed Alpha によって知りうること(Trusted Tester Agreement!)

前回のあらすじ

前回のあらすじ

Java 8 ランタイム以降のサンドボックスとgVisor より

- やっぱり AppEngine ja night #3 で GAE/SE と gVisor の関係について話した

前回のあらすじ - GAE 文脈の gVisor

GAE/Java 7 サンドボックス

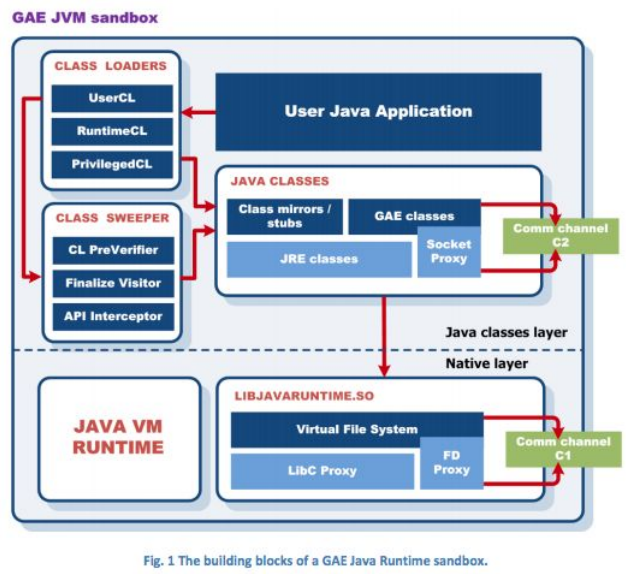
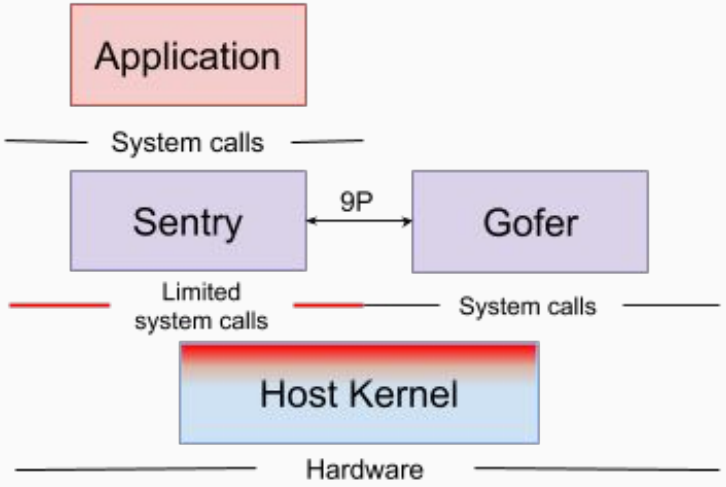


Fig. 1 The building blocks of a GAE Java Runtime sandbox.

SE-2014-02 よりGAE/Java 7 サンドボックス

移行

GAE/Java 8 ～ サンドボックス(の一部)



google/gvisor より gVisor アーキテクチャ

従来の制限が強くランタイムごとに実装していたサンドボックスを、リソース利用効率を損なわずにより汎用的かつ自由に実現するもの

前回のあらすじ - まとめ

- 現在 GAE/SE は gVisor ベースのサンドボックスへの移行過渡期
  - より自由に
  - より汎用的に
- 続々と集まる GAE/SE ランタイムたち
  - Java 8 (2017/9 GA)
  - Node.js 8 (2018/6 Beta)
  - Python 3 (EAP)
  - PHP 7 (EAP)
  - (続く)
- 冷遇時代もあったけど GAE/SE の未来は明るい
- ハジコン

GAE 以外は？

Cloud Functions は gVisor ベース

- gVisor の runsc と同様の /proc が見えることを s.hiruta が発見
- /proc の中身が一通り gVisor のドキュメントと一致することを確認
  - gist:apstndb/GCF_GVISOR_EVAL.md
- 開発者から gVisor のイースターエッグのネタバレ

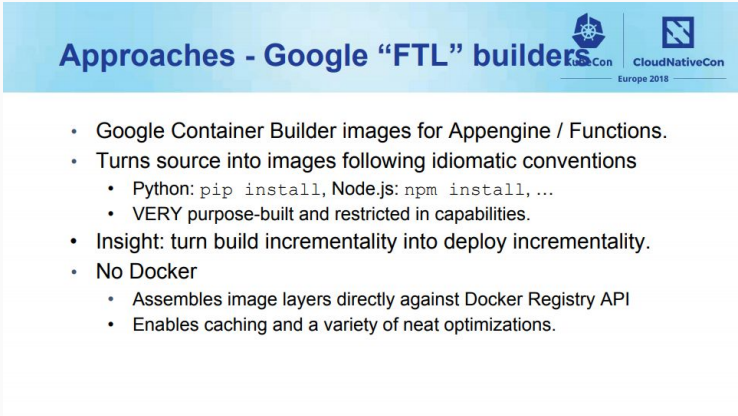
GCPUG Slack #g-c-functions_ja より

Cloud Functions の展望

- サンドボックスは GAE 同様 gVisor ベース
- ランタイムのバージョンアップや対応増に期待
  - Node.js 6 (2017/3 beta)
  - Node.js 8 (EAP)
  - Python 3 (EAP)
    - googlecloud-community Slack(NDA 不要)等で募集
  - (続く)

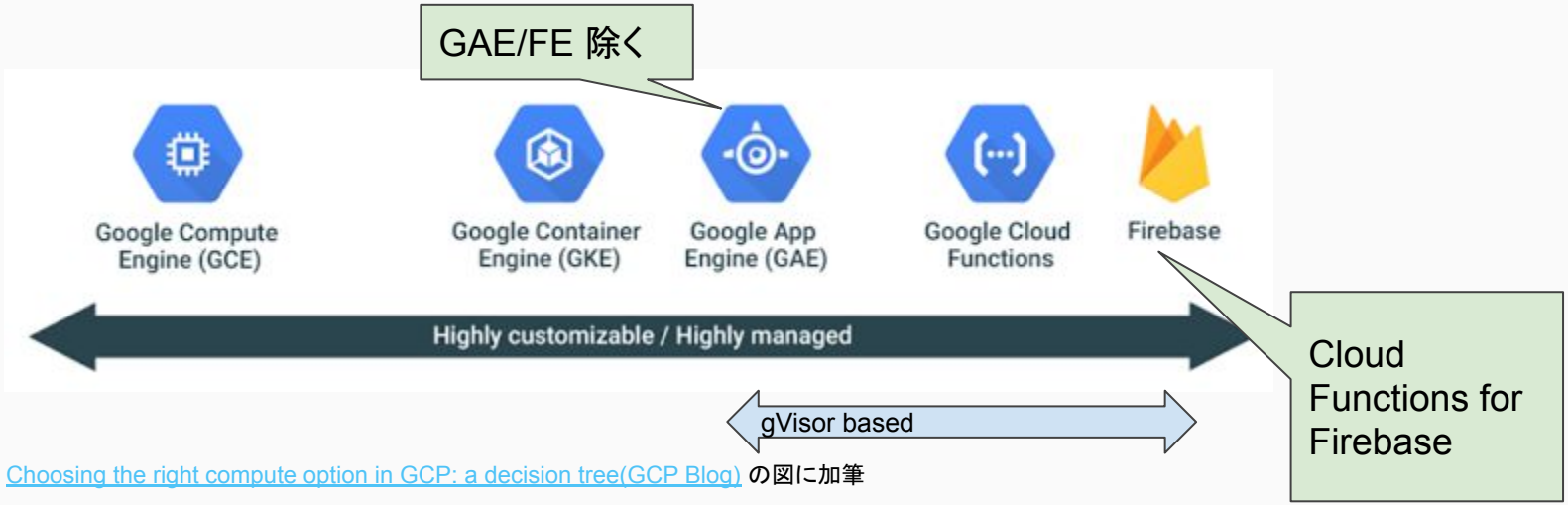
Cloud Functions の展望 - FTL

- ビルドパイプラインも GAE と同様 FTL になりそう
  - FTL Design Docs とか Going from Source to Image



Going from Source to Image より

Compute Options と gVisor



Choosing the right compute option in GCP: a decision tree(GCP Blog) の図に加筆

- Compute Options の半分が gVisor ベースに
- gVisor は Google におけるサーバーレス戦略と密接に関係

共通するキーパーソン

- gVisor とプロダクトの設計に深く関わっている人物が存在
  - App Engine
  - Cloud Functions
  - Cloud ML Engine

KubeCon 2018 EU Container Isolation at Scale (Introducing gVisor) より

Zhengyu He, Staff Software Engineer and Manager of Google Cloud, Ph.D., Georgia Tech. Co-founder and Tech Lead of gVisor. Led the design of multiple Google Cloud products including App Engine, Cloud Functions, and Cloud ML.

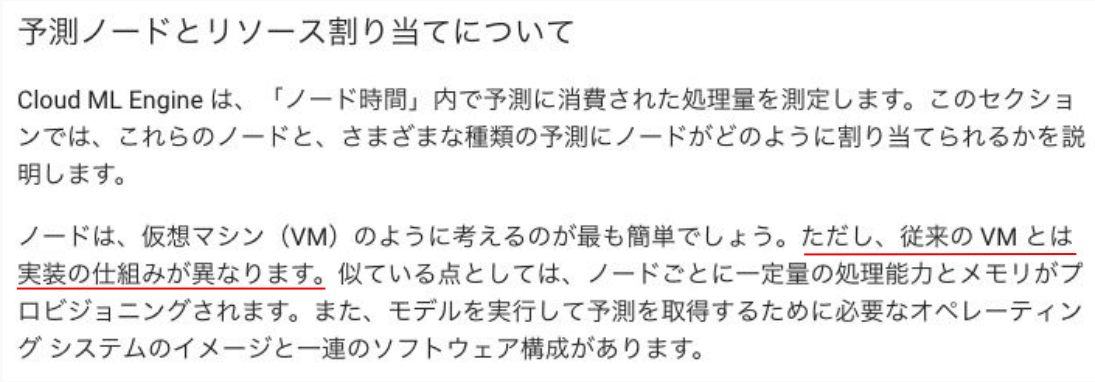
http://www.tech-meetup.com/bayarea/2018_06_03_gvisor_alluxio より

GAE/Node.js 公開直後に情報開示

- Director Product Manager より Cloud Functions & Cloud ML Engine でも gVisor ベースの技術を使っていることがはじめて明言
- Cloud ML Engine は Online Prediction で使っているらしい？(未検証)

Cloud ML Engine の Online Prediction とは？

- デプロイした学習済モデルを推論 API として提供する機能
  - TensorFlow だけでなく XGBoost, scikit-learn を選択可能
    - サードパーティなのでサンドボックスしたい
- 実行/課金単位はノード(ただし VM ではないと明記)
  - GAE のインスタンス相当？
- GAE と同様に gVisor を利用する理由がある



予測ノードとリソース割り当てについて

Cloud ML Engine は、「ノード時間」内で予測に消費された処理量を測定します。このセクションでは、これらのノードと、さまざまな種類の予測にノードがどのように割り当てられるかを説明します。

ノードは、仮想マシン（VM）のように考えるのが最も簡単でしょう。ただし、従来の VM とは実装の仕組みが異なります。似ている点としては、ノードごとに一定量の処理能力とメモリがプロビジョニングされます。また、モデルを実行して予測を取得するために必要なオペレーティング システムのイメージと一連のソフトウェア構成があります。

Cloud ML Engine ドキュメントより

まとめ

- App Engine 以外も Google が完全に信用できるもの以外は gVisor でサンドボックスする流れ
  - Cloud Functions と Cloud ML Engine は明言

- 今 VM ベースなものも移行しないだろうか？
  - Cloud Memorystore とか Cloud Composer とか…

残る謎

- gVisor の動作確認リストの裏にあるロードマップは
- gVisor の checkpoint/restore が runsc からも使えるようになったが、どこで使われるか

What works?

The following applications/images have been tested:

- golang
- httpd
- java8
- jenkins
- mariadb
- memcached
- mongo
- mysql
- nginx
- node
- php
- postgres
- prometheus
- python
- redis
- registry
- tomcat
- wordpress

google/gvisor より

参考情報

- gVisor in depth
  - DockerCon’18 で Google の中の人が大体話したという話