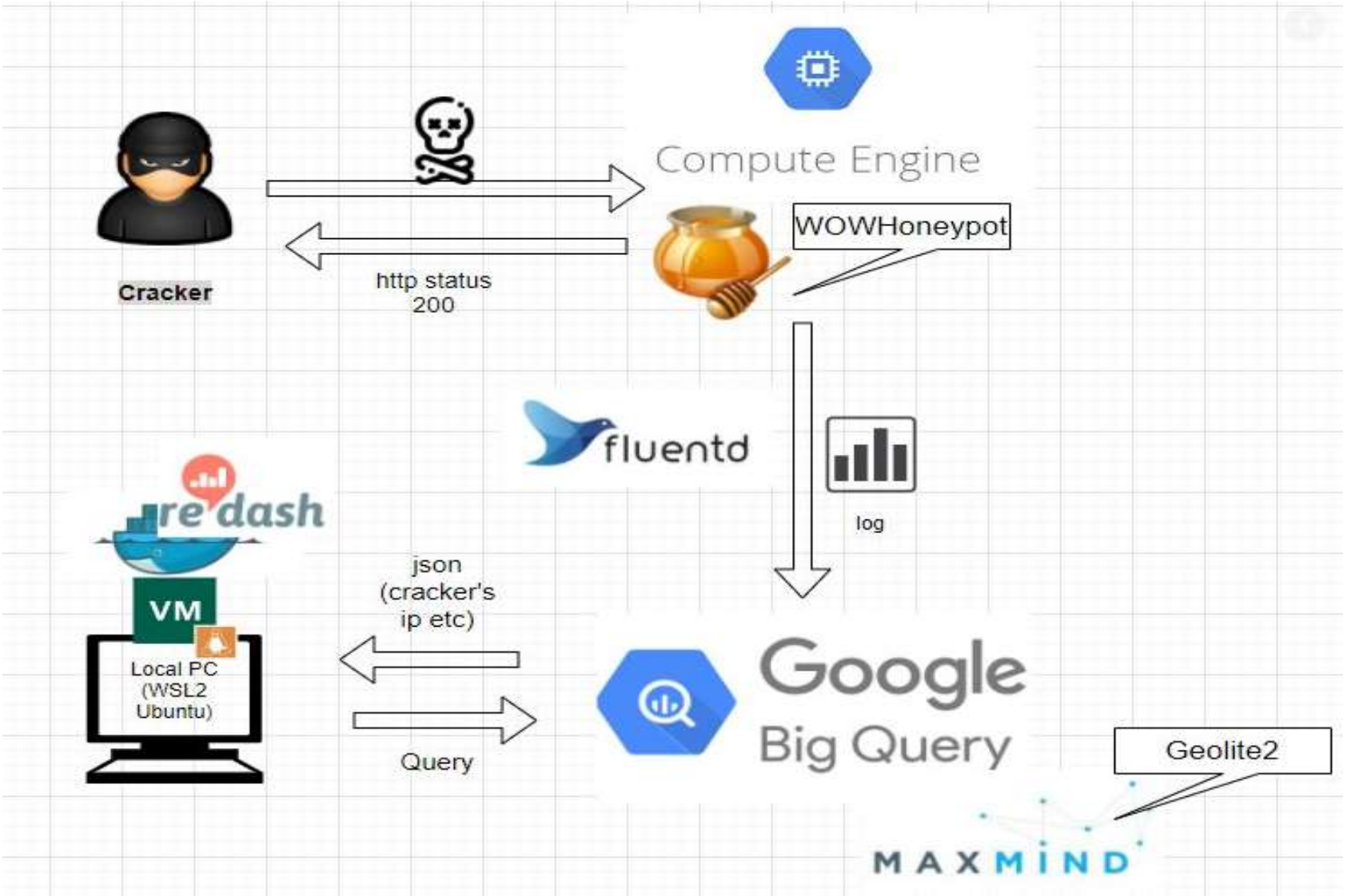
ハニーポッターと
謎のウィルス
~ハニーポット構築過程のご紹介~

本物のウィルスに
会いに行こう！！

2020/10/12

関西事業所　中田健太

概要


Compute Engine
WOWHoneypot
Cracker
http status
200
fluentd
log
re dash
VM
Local PC
(WSL2
Ubuntu)
json
(cracker's
ip etc)
Query
Google
Big Query
Geolite2
MAXMIND

ハニーポットとは

・ネットワーク上で攻撃者やウイルス
などをおびき寄せるために
わざと攻撃しやすいように見せかけた
コンピュータなど

・**低対話型と高対話型**

・今回は**WOWHoneypot**を使用
→主に右の本を参照

Google Cloud

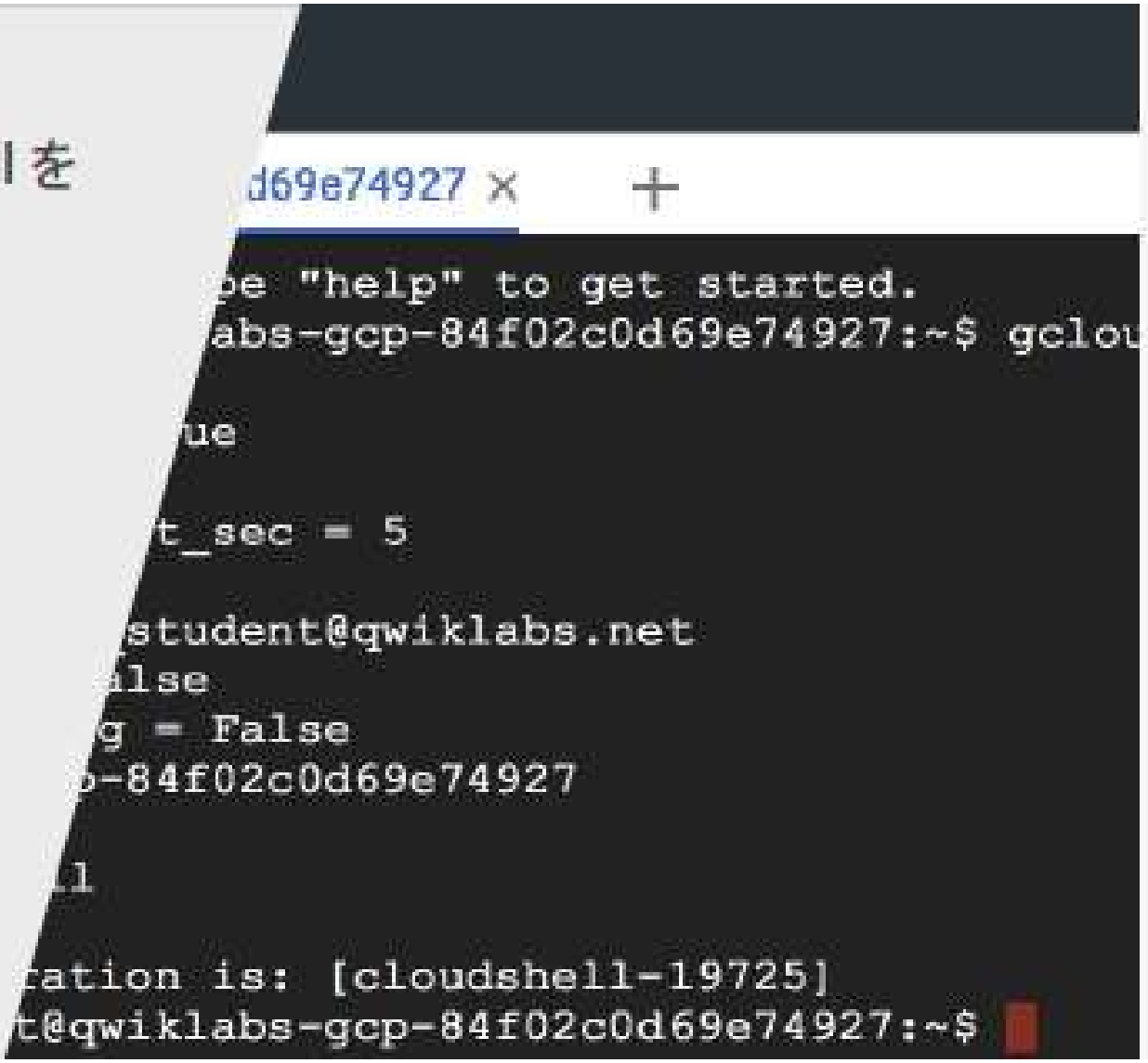
Google Compute Engine (GCE)

- 高性能でスケーラブルな Virtual Machine (VM)
  - CPU / メモリを柔軟に構成可能
  - ディスクを用途に合わせて選択可能
    - 標準 HDD、SSD、ローカル SSD
- 低コスト、自動割引
  - 課金は最低 1 分、その後秒単位での課金
  - 事前申込不要で継続利用割引が適用
- 透過的なメンテナンス
  - ライブマイグレーションにより、
    メンテナンス時に稼働中の VM が自動的に移動し、
    VM の再起動が不要
- 後述する GCP のネットワーキング機能との統合

Google Cloud

デモ：Cloud SDK / Cloud Shell を使ってみよう

- Cloud Console から Cloud Shell を開く
- Cloud SDK によって提供される、gcloud・gsutil・ bq 等のコマンドを実行
- Cloud Shell は Cloud SDK や Git 等の各種ツールがインストール済みの VM で、無料利用



Google Cloud

Virtual Private Cloud (VPC)

- Virtual Private Cloud (VPC) によるリソースの相互接続や分離
  - リージョンを跨いだ VPC ネットワークが構築可能
    - 同一ネットワーク内では内部 IP で通信可能
  - 内部 IP アドレス範囲を自由に設定
  - 経路やファイヤーウォールを柔軟に設定
- Cloud VPN
  - 既存のネットワークと GCE が存在する VPC ネットワークを IPsec 経由で安全に接続
- その他様々なネットワークサービス
  - Cloud Load Balancing
  - Cloud CDN
  - Cloud Interconnect　　etc.,

Google Cloud

BigQuery

- フルマネージドの No-Ops データウェアハウス
- ペタバイト規模で高速
- スタンダード SQL
- 耐久性があり高可用性を備える

BigQuery は 2 つのサービスが一体化されたもの

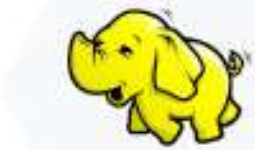
Google
BigQuery

- 高速な SQL クエリエンジン
- データセット用のマネージド ストレージ

Google Cloud

ビッグデータ対応のリッチな オープンソース エコシステム

http://hadoop.apache.org/
http://pig.apache.org/
http://hive.apache.org/
http://spark.apache.org/



1 Hadoop は標準的な オープンソースの MapReduce フレームワーク

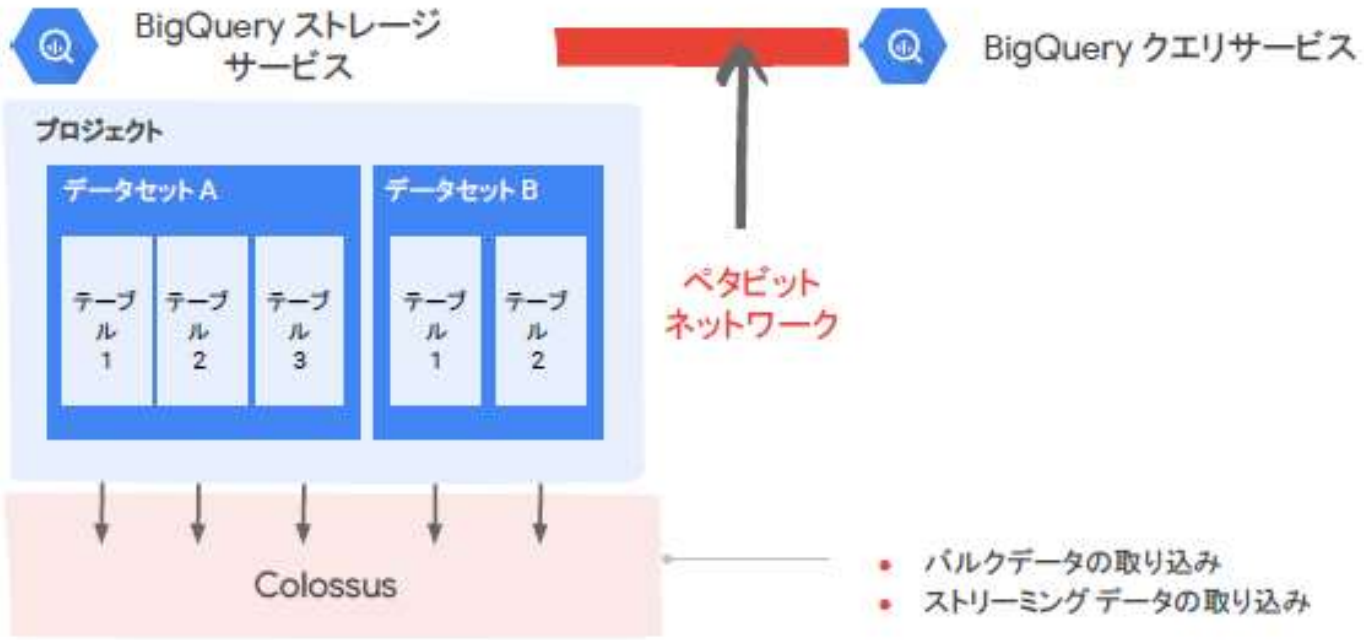
2 Pig は便利な スクリプティング言語で、 Hadoop MapReduce ジョブに コンパイル可能

3 Hive は データ ウェアハウス システムおよびクエリ言語

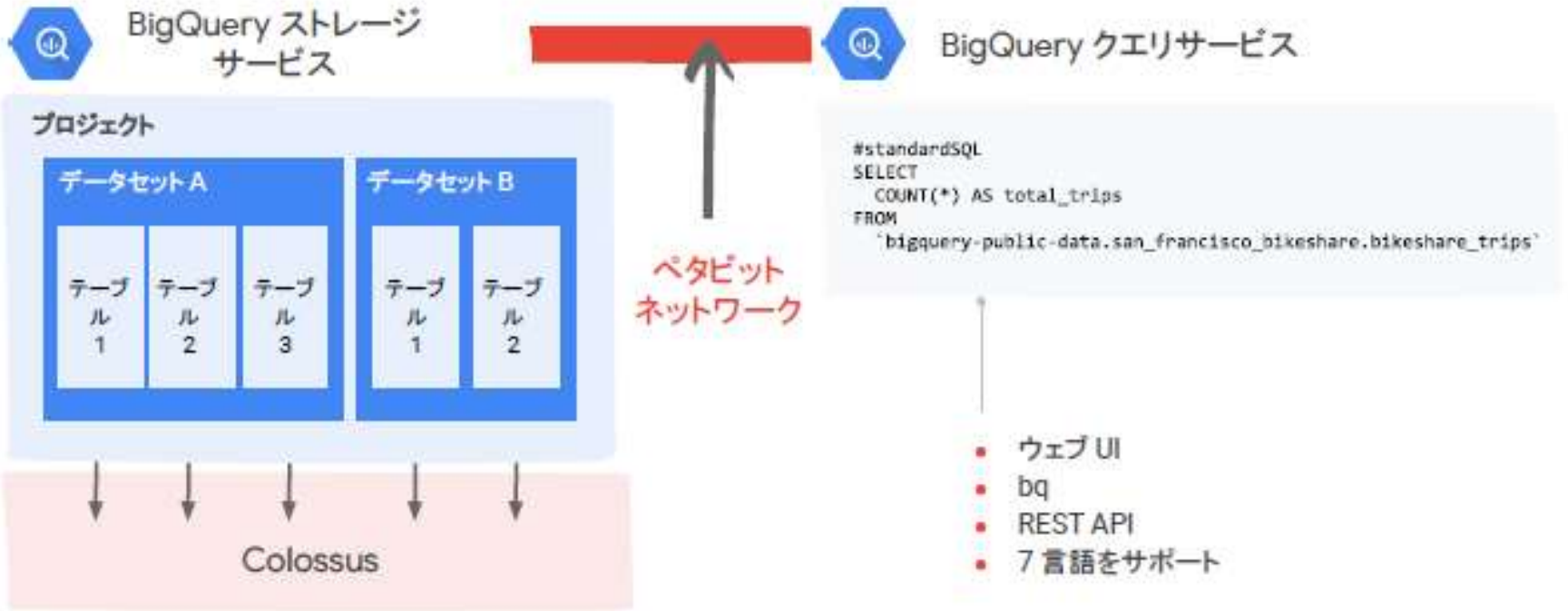
4 Spark は、SQL、ストリーミング、 機械学習に使用できる高速で インタラクティブな 汎用フレームワーク

Google Cloud

BigQuery の仕組み

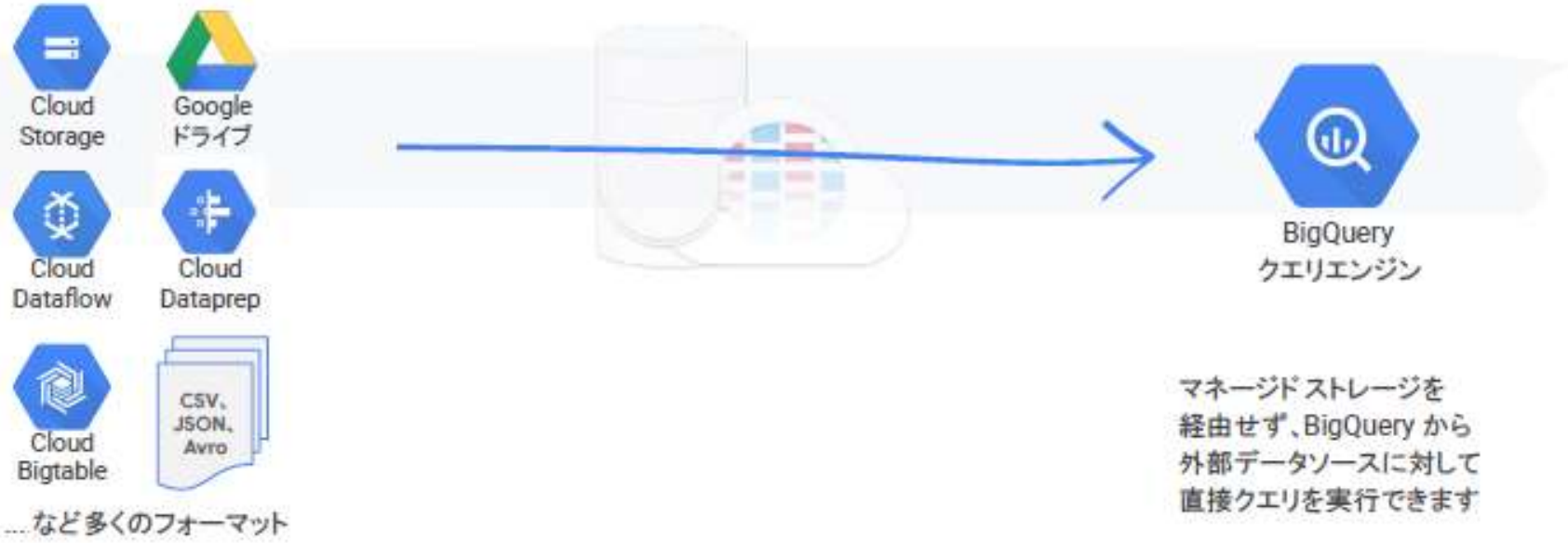


BigQuery の仕組み



Google Cloud

BigQuery は、GCS やドライブで、外部（フェデレーション）データソースを直接照会できる



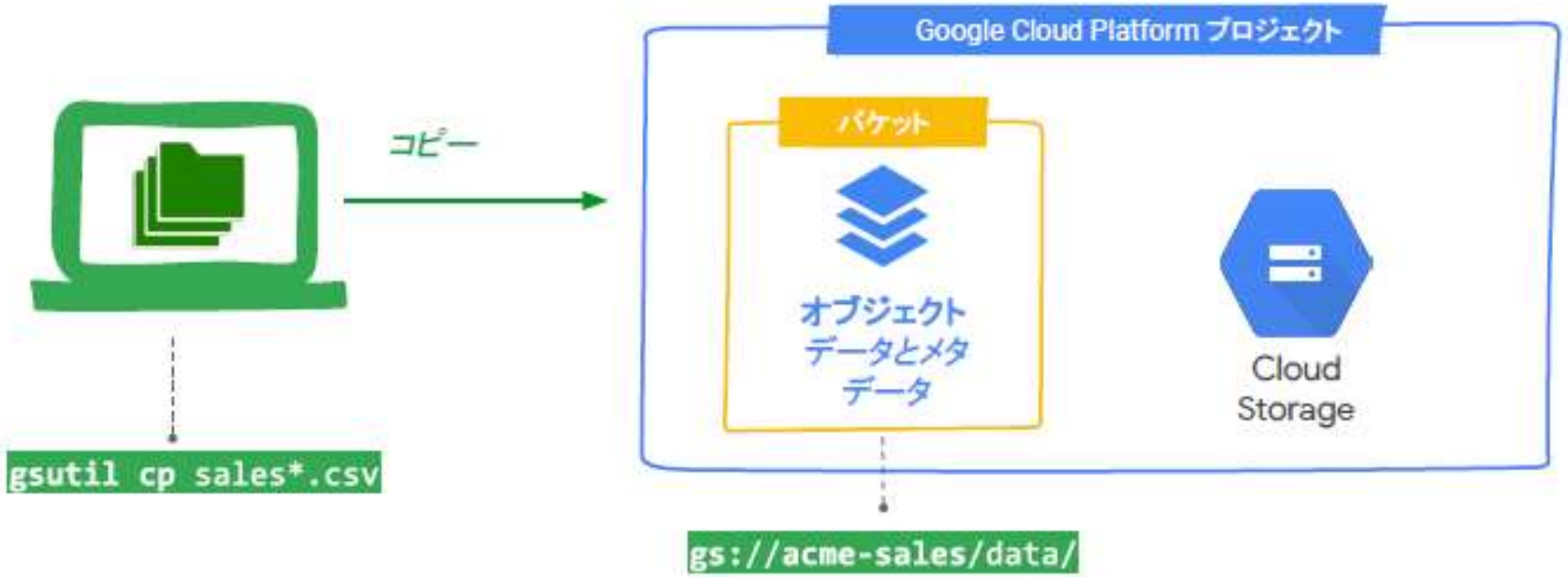
Google Cloud

Google の 99.999999999% の耐久性を持つストレージを利用

Cloud Storage

Google Cloud

データがすでにあるなら、gsutil ツールで速やかにデータをクラウドに移行できる



Google Cloud

gsutil = Google Storage Utility, cp = コピー

・**OSSのデータログ収集ツール**

・**リアルタイム**に生成されるデータ
に対しては有効なツール（↔emblulk)

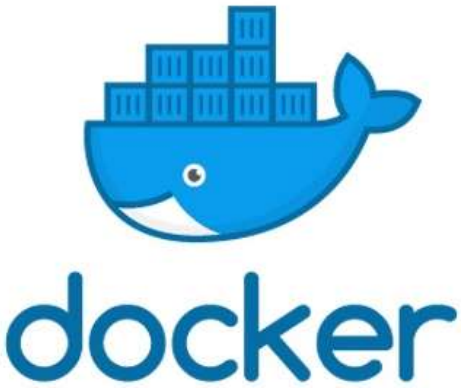
・プラグインによって**データの加工まで実施**
（access_logがBASE64だったので、今回はbase64デコードを利用）

・OSS で提供されている
ダッシュボード生成ツール

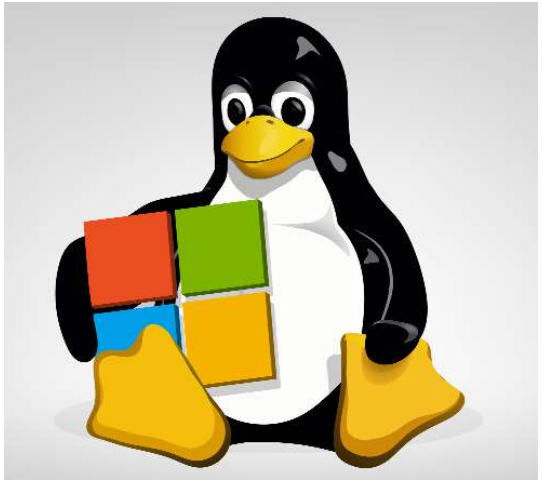
・データソースとして
MySQL や PostgreSQL 等の RDMS 、
Amazon RDS の Aurora や、Elasticsearch、
Google BigQuery, Graphite を指定可

```
SELECT
    IFNULL(country_name, 'Other') AS country_name,
         SUM(1)
FROM (
  SELECT
    clientip,
    NET.IPV4_TO_INT64(NET.IP_FROM_STRING(clientip)) AS clientIpNum,
    TRUNC(NET.IPV4_TO_INT64(NET.IP_FROM_STRING(clientip))/(256*256)) AS classB
  FROM
    `WoWHoneypotのテーブル名（ワイルドカード指定可能）` ) AS a
LEFT OUTER JOIN
  `geolite2_cityのテーブル名` AS b
ON
  a.classB = b.classB
  AND a.clientIpNum BETWEEN b.startIpNum AND b.endIpNum
GROUP  BY  country_name
```

Docker for windows と WSL2



・必要最低限のものが入っており、
まとめておけばお手軽に環境を
用意可
(Re;dashの環境を用意するために
使用)



・WindowsでLinuxをそのまま
つかえるようにしたもの

今回の構築に関する縛りルール

・90%コマンドのみで構築
　- VPCとかもすべて！

・可視化までする

実際のシステムを見てみましょう！