Google Cloud Platform Solutions for DevOps Engineers

Márton Kodok / @martonkodok
Google Developer Expert at REEA.net - Targu Mures

March 2019 - Vilnius, Lithuania
About me

- Among the Top3 romanians on Stackoverflow 130k reputation
- Google Developer Expert on Cloud technologies
- Crafting Web/Mobile backends at REEA.net
- BigQuery/Redis and database engine expert
- Active in mentoring and IT community

StackOverflow: pentium10
GitHub: pentium10
Slideshare: martonkodok
Twitter: @martonkodok
1. Application development in the Cloud
2. **App Engine**: Scale your apps seamlessly from zero to planet scale
3. Cloud **Functions**: your gateway to GCP Services
4. Google **Stackdriver**: Metrics, logging, alerting are a universal right!
5. **BigQuery**: federated data access warehouse
6. Doing **hybrid cloud** mixing on premise with cloud
7. Practical use cases
8. **Qwiklabs**
What’s so hard about traditional app development?
REEA.net uses GCP

Build on the same infrastructure that powers Google
Google sees serverless as

<table>
<thead>
<tr>
<th>Programming model</th>
<th>Operational model</th>
<th>Billing model</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔ Focus on code</td>
<td>✔ Zero ops</td>
<td>✔ Pay for usage</td>
</tr>
<tr>
<td>✔ Event-driven</td>
<td>✔ Automatic scaling</td>
<td></td>
</tr>
<tr>
<td>✔ Stateless</td>
<td>✔ Managed security</td>
<td></td>
</tr>
</tbody>
</table>

Dev

Ops

$
Serverless is more than a set of functions
Meet Serverless

- Automatic scaling
- No server management
- Only pay for what you use
- Event-driven

NOT

serverless data center depicted
Serverless is about maximizing **elasticity**, **cost savings**, and **agility** of cloud computing.
Serverless types

Platforms | Triggered Code
**App Engine - managed application platform**

- **Google Cloud Platform**
  - API 1
    - App Engine
    - Split App Versions
  - API 2
    - Compute Engine
    - Virtual Machines
  - API 3
    - Container Engine
    - Kubernetes Services

- **Cloud Load Balancing**

**App Engine**

- Scale your applications seamlessly from zero to planet scale
- Automatically scales depending on your application traffic
- Traffic Splitting (app versions, A/B tests, incremental rollouts)

**Best used:**
- For HTTP services
- For existing applications
App Engine: Traffic splitting

A/B testing and Canary releasing with a few clicks or a single gcloud command
App Engine: Runtimes

- Node.js
- Python
- PHP 7.2
- Java 8
- Go 1.11
Async message processing
Asynchronous task execution

Cloud Tasks
Message queue system
Guaranteed at-least-once delivery
Future Scheduling
Will only be dispatched once on dups

Best used:
- For HTTP services
- For App Engine handlers

Cloud PubSub
Deliver event data based on topics
Exactly-once processing
Build multi-cloud on premise, hybrid apps
Cross zone message replication

Best used:
- For large-scale ingestion of events, streams
- Topics, publish/subscribe patterns, IoT
Cloud Tasks - Message Queue systems
**Task Queues**

---

**Push queues**
- Put with delays

**Pull queues**
- Ability to “tag”
- Lease multiple eg: gameboard updated, game id as tag.

---

**Slides:** [https://www.slideshare.net/martonkodok](https://www.slideshare.net/martonkodok)

**Title:** Architectural Patterns - Message Queues
Reduce request latency

Front-end response times (ms)
daily aggregate

Time (ms)

0 250 500 750 1000 1250 1500 1750 2000 2250

Mon Tue Wed Thu Fri Sat Sun

FE-1 FE-2 FE-3 FE-4

Highcharts.com

GCP Solutions for DevOps Engineers
@martonkodok
Cloud Functions
Cloud Functions

**Unit:** Function
**Trigger:** Events and HTTP

**Best used:**
- For Events & Async workloads
- For single-purpose microservices
Cloud Functions - Event-driven - FaaS

- Simplest way to run your code in the cloud - just deploy
- Automatically scales, highly available and fault tolerant
- No servers to provision, manage, patch or update
- Pay only while your code runs
- Connects and extends cloud services

(In alpha: Java)

Node 8.11
- Support for async/await

Python 3.7.1
- Flask microframework

Go 1.11
- Familiar building blocks
Functions: your gateway to GCP Services

Access 20+ Google services from GCF

- BigQuery
- Cloud Pub/Sub
- Cloud Storage
- Cloud Bigtable
- Cloud Datastore
- Cloud Machine Learning
- Cloud Vision API
- Cloud Speech API
- Cloud Spanner
- Cloud Natural Language API
- Cloud Translation API
- Stackdriver
- Firebase
- Assistant

GCP Solutions for DevOps Engineers

@martonkodok
Cloud Functions Trigger

**Firebase**
- Realtime DB change
- Authentication
- Remote config
- Google Analytics log

**BigQuery**
- Job create
- Job complete

**Cloud Firestore**
- Create
- Update
- Write
- Delete
<table>
<thead>
<tr>
<th><strong>Summary: When to use what</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloud Functions</strong></td>
</tr>
<tr>
<td>Smallest unit of computing</td>
</tr>
<tr>
<td>Event driven architecture</td>
</tr>
<tr>
<td>Connect &amp; extend services</td>
</tr>
</tbody>
</table>

GCP Solutions for DevOps Engineers

@martonkodok
Slice and dice your application any way you think it makes sense to be more agile, to build better apps that deliver value.
How I Learned How to Stop Worrying and Dig Hosted Monitoring
What to do for monitoring?

- In-house or hosted?
- Modern vs proven?
- Specialized or general?
- Paid vs free?

Let’s be honest:
- Not enough time
- Small environment
- More important work to do
- Everyone else is doing it
Google Stackdriver

Intelligent monitoring and management for services, containers, applications, and infrastructure.

- Works with GCP, AWS, on prem
- Logging, tracing, alerts
- Collectd agent
- Cost per resource
- Easy point and click alerting policy
- Debugger, Error reporting, profiler *free
- Uptime checks *free
Google Stackdriver: Uptime check (offered free)

- Check Type: HTTP, TCP
- Resource Type: URL, Instance, App Engine, Load Balancer
- Global Locations
- Custom Headers (encrypted)
- Authentication
- Whitelisted source IPs
- Alerting policy: Email, SMS, Slack, PagerDuty, Hipchat, Campfire, Webhooks
Google Stackdriver: Uptime check

Uptime Check latency

<table>
<thead>
<tr>
<th>checker_location</th>
<th>instance_id</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pac-singapore</td>
<td></td>
<td>1.703s</td>
</tr>
<tr>
<td>eur-belgium</td>
<td></td>
<td>433.23ms</td>
</tr>
<tr>
<td>sa-brazil-sao_paulo</td>
<td></td>
<td>1118.96ms</td>
</tr>
<tr>
<td>usa-iowa</td>
<td></td>
<td>1024.70ms</td>
</tr>
</tbody>
</table>
Life of a log

- Log entry
- Logging API
  - Export?
  - Exclude?
    - YES: Log entry discarded
    - NO: Log entry ingested
      - Billing is calculated based on Log Entries ingested
- Destination selector
  - Cloud Storage
  - BigQuery
  - Cloud Pub/Sub
Metrics, logging, alerting are a universal right!
What is BigQuery?

Analytics-as-a-Service - Data Warehouse in the Cloud

Scales into Petabytes on Managed Infrastructure - load up to 5TB large files

SQL 2011 + Javascript UDF (User Defined Functions)

Familiar DB Structure (table, columns, views, struct, nested, JSON)

Integrates with Google Sheets + Cloud Storage + Pub/Sub connectors

Decent pricing (storage: $20/TB  cold: $10/TB, queries $5/TB)  *March 2019
Architecting for The Cloud

On-Premises Servers

- Frontend Platform Services
- Event Sourcing
- Metrics / Logs / Streaming

Pipelines

- ETL Engine

Google Cloud Platform

BigQuery

GCP Solutions for DevOps Engineers

@martonkodok
BigQuery: federated data access warehouse

Audit logs
Billing entries
Stackdriver
Cloud Dataflow
Database SQL
Real-Time Events
Multiple Platforms
Frontend Platform Services

Application & Presentation
Analysis
Processing
ML
BI Interface
Data Studio 360

Report & Share
Business Analysis

GCP Solutions for DevOps Engineers
@martonkodok
Data needs to be processed in multiple services. How can we pipe to multiple places?
We have our app outside of GCP. How can we use the benefits of BigQuery?
Data Pipeline Integration at REEA.net

On-Premises Servers
- Frontend Platform Services
- Event Sourcing
- Metrics / Logs / Streaming

Pipelines
- FluentD

Application
- Servers

Database
- SQL

Google Cloud Platform
- Cloud Storage
- Load / Export Replay
- Cloud Functions
- Analytics Backend BigQuery

Report & Share Business Analysis
- Tableau
- QlikView
- Data Studio
- Internal Dashboard

Development Team

Standard Devices HTTPS

GCP Solutions for DevOps Engineers

@martonkodok
The following slides will present a sample Fluentd configuration to:

1. **Transform** a record
2. **Copy** event to multiple outputs
3. **Store** event data in File (for backup/log purposes)
4. **Stream** to BigQuery (for immediate analyses)
Filter plugin mutates incoming data. Add/modify/delete event data **transform attributes without a code deploy**.

The copy output plugin copies events to multiple outputs. File(s), multiple databases, DB engines. **Great to ship same event to multiple subsystems.**

The Bigquery output plugin on the fly streams the event to the BigQuery warehouse. **No need to write integration.** Data is available immediately for querying.

Whenever needed other output plugins can be wired in: Kafka, Google Cloud Storage output plugin.
<filter frontend.user.>*>
   @type record_transformer
   enable_ruby
   remove_keys host
   <record>
      bq {"insert_id":"${uid}","host":"${host}",
      "created":"${time.to_i}"}
      avg ${record["total"] / record["count"]}
   </record>
</filter>

syntax: Ruby, easy to use.

Great for:
- date transformation,
- quick normalizations,
- calculating something on the fly, and store in clear log-analytics db
- renaming without code deploy.

record_transformer copy file BigQuery
<match frontend.user.*>(
    @type copy
    <store>
        @type forest
        subtype file
        <template>
            path /tank/storage/${tag}.*.log
            time_slice_format %Y%m%d
        </template>
    </store>
</match>
<match frontend.user.*>
  @type bigquery
  method insert
  auth_method json_key
  json_key /etc/td-agent/keys/key-31da042be48c.json
  time_field timestamp
  time_slice_format %Y%m%d
  table user$%{time_slice}
  ignore_unknown_values
  schema_path /etc/td-agent/schema/user_login.json
</match>

Connector uses:
- JSON key auth file
- JSON table schema

Pro features:
- streaming to Partitioned tables
- ignore unknown values
  (not reflected in schema)
Our benefits using BigQuery

- SQL+JSON (big costs saving with partitioning/clustering)
- run raw ad-hoc queries (either by analysts/sales or Devs)
- ability to throw in / join all kind of data
- pricing model 1TB free every month
- no more throwing away-, expiring-, aggregating old data
- no provisioning/deploy
- no running out of resources
- no more focus on large scale execution plan
# Qwiklabs - Find Your Topic and Skill Level

<table>
<thead>
<tr>
<th>Topic</th>
<th>Kubernetes</th>
<th>Machine Learning</th>
<th>Certification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introductory</strong></td>
<td><strong>Baselines: Infrastructure Quest</strong>&lt;br&gt; 1. Kubernetes Engine: Qwik Start (lab)&lt;br&gt; (video) - Deploy a containerized application with Kubernetes Engine.&lt;br&gt; 2. IoT: Qwik Start (lab) - Learn about GCP’s Cloud IoT Core service, where almost any IoT thing is possible!</td>
<td><strong>Baselines: Data, ML, AI Quest</strong>&lt;br&gt; 1. Cloud ML Engine: Qwik Start (lab)&lt;br&gt; (video) - Train and deploy a TensorFlow model to Cloud ML Engine.&lt;br&gt; 2. Cloud Natural Language API: Qwik Start (lab) - Perform sentiment analysis on a block of text. For fun, paste in the last email you sent!</td>
<td><strong>Cloud Architecture Quest</strong>&lt;br&gt; 1. Stackdriver Qwik Start: Web UI (lab) - What’s up with your VM? Learn to install monitoring and logging agents to collect info about your GCP resources.&lt;br&gt; 2. Getting Started with Cloud KMS (lab) - Create your first KeyRing and CryptoKey with GCP.</td>
</tr>
<tr>
<td><strong>Experienced</strong></td>
<td><strong>Baselines: Kubernetes Quest</strong>&lt;br&gt; 1. Kubernetes Engine: Qwik Start (lab)&lt;br&gt; (video) - Deploy a containerized application with Kubernetes Engine.&lt;br&gt; 2. Intro to Docker (lab) - Are your containers running? Create, run, and debug containers, then pull and push images to and from Google Container Registry.</td>
<td><strong>Baselines: Machine Learning APIs Quest</strong>&lt;br&gt; 1. Cloud ML Engine: Qwik Start (lab)&lt;br&gt; (video) - Train and deploy a TensorFlow model to Cloud ML Engine.&lt;br&gt; 2. Detect Labels, Faces, and Landmarks in Images with the Cloud Vision API (lab) - How do you feel about self-driving cars? Learn what’s behind the AI making the future possible.</td>
<td><strong>Cloud Architecture Quest</strong>&lt;br&gt; 1. Stackdriver Qwik Start: Web UI (lab) - What’s up with your VM? Learn to install monitoring and logging agents to collect info about your GCP resources.&lt;br&gt; 2. Application Performance Management (APM) with Stackdriver (lab) - Dig deeper into Stackdriver with a problematic environment - can you identify the issues?</td>
</tr>
</tbody>
</table>

---

**GCP Solutions for DevOps Engineers**

@martonkodok
Thank you.

Slides available on: slideshare.net/martonkodok

Reea.net - Integrated web solutions driven by creativity to deliver projects.