A Precise Model for Google Cloud Platform

Stéphanie Challita | Faiez Zalila | Christophe Gourdin | Philippe Merle
Inria Lille – Nord Europe & University of Lille
6th IEEE International Conference on Cloud Engineering (IC2E 2018)
Google Cloud Platform

Google Cloud Platform

Coca-Cola
AIRBUS
intel
HTC
Spotify
redhat
EQUINIX
BNP PARIBAS
MOTOROLA

17 – 20 April, 2018, Orlando, Florida, USA
An agreement with the developer on exactly how the system will operate

GCP documentation is written in natural language → human errors and/or semantic confusions
Agenda

- Drawbacks & Motivations
- Contribution
- Perspectives
Drawbacks & Motivations
List of Drawbacks

- Informal Heterogeneous Documentation
- Imprecise Types
- Implicit Attribute Metadata
- Hidden Links
- Redundancy
- Lack of Visual Support
### Property name

<table>
<thead>
<tr>
<th>Property name</th>
<th>Value</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPv4Range</td>
<td>string</td>
<td>The range of internal addresses that are legal on this network. This range is a CIDR specification, for example: 192.168.0.0/16. Provided by the client when the network is created.</td>
<td></td>
</tr>
</tbody>
</table>

Available at
https://cloud.google.com/compute/docs/reference/latest/networks

### Vs.

### Fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>projectId</td>
<td>string</td>
<td>Required. The Google Cloud Platform project ID that the cluster belongs to.</td>
</tr>
</tbody>
</table>

Available at
https://cloud.google.com/dataproc/docs/reference/rest/v1/projects.regions.clusters
### Imprecise Types

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>selfLink</td>
<td>string</td>
<td>[Output Only] Server-defined URL for the resource.</td>
</tr>
<tr>
<td>email</td>
<td>string</td>
<td>The email address of the service account. Note: This field is used in responses only. Any value specified here in a request is ignored.</td>
</tr>
</tbody>
</table>
| instanceClass| string     | Instance class that is used to run this version. Valid values are:  
- AutomaticScaling: F1, F2, F4, F4_1G  
- ManualScaling or BasicScaling: B1, B2, B4, B8, B4_1G  
Defaults to F1 for AutomaticScaling and B1 for ManualScaling or BasicScaling. |
| locations[]   | string     | The list of Google Compute Engine locations in which the cluster's nodes should be located. |

Available at:
- `https://cloud.google.com/compute/docs/reference/latest/targetHttpsProxies`
- `https://cloud.google.com/iam/reference/rest/v1/projects.serviceAccounts`
- `https://cloud.google.com/appengine/docs/admin-api/reference/rest/v1beta5/apps.services.versions`
Lack of Visual Support

- Only **descriptive information** in GCP documentation
  - A huge time to be properly understood and analyzed

- Visual diagrams easily highlight in short but catchy view the concepts of the API
  - Help to avoid wastage of time
  - Logical sequence and comparative analysis can be undertaken
Contribution
Global Vision

- Precise model for GCP
- Work of reverse-engineering, HTML Model
- GCP model refinement
GCP Snapshot

- GCP engineers could update/correct GCP documentation
- Continuously following up with GCP documentation is costly
- Snapshot of GCP API
GCP Crawler & GCP Model

- **GCP Crawler** to extract all GCP resources, their attributes and actions
- **GCP Model** for a better description of the GCP resources and for reasoning over them

No more Informal Heterogeneous Documentation

GCP Crawler & GCP Model

- GCP Crawler to extract all GCP resources, their attributes and actions
- GCP Model for a better description of the GCP resources and for reasoning over them

No more Informal Heterogeneous Documentation

GCP Crawler & GCP Model

- GCP Crawler to extract all GCP resources, their attributes and actions
- GCP Model for a better description of the GCP resources and for reasoning over them

No more Informal Heterogeneous Documentation
Type Refinement

- By adopting the data type system proposed by OCCLware metamodel
  - defining regular expressions,
  - and using the EMF validator to check the type constraints that are attached to the attributes

- If the type of an attribute in the documentation is string and the description explains that this is an email address, we apply the email validation constraint:
  - STRINGTYPE + this regular expression:
    \^[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,6}$

- No more Imprecise Types 😊
Model Visualization

For an easier understanding and analysis of the API, because the insights become obvious.

No more Lack of Visual Support 😊
“GCP documentation is developed by two separate clusters of development teams”

“GCP documentation is redundant and not comprehensive”

<table>
<thead>
<tr>
<th>Redundant Attributes</th>
<th>Before Abstraction</th>
<th>After Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>name</td>
<td>92</td>
<td>64.79%</td>
</tr>
<tr>
<td>id</td>
<td>80</td>
<td>56.34%</td>
</tr>
<tr>
<td>selfLink</td>
<td>79</td>
<td>55.63%</td>
</tr>
<tr>
<td>kind</td>
<td>79</td>
<td>55.63%</td>
</tr>
<tr>
<td>description</td>
<td>75</td>
<td>52.82%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>57.04%</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Redundant Actions</th>
<th>Before Abstraction</th>
<th>After Abstraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>%</td>
</tr>
<tr>
<td>get</td>
<td>142</td>
<td>100.00%</td>
</tr>
<tr>
<td>list</td>
<td>142</td>
<td>100.00%</td>
</tr>
<tr>
<td>delete</td>
<td>140</td>
<td>98.59%</td>
</tr>
<tr>
<td>insert</td>
<td>76</td>
<td>53.52%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>88.02%</strong></td>
<td></td>
</tr>
</tbody>
</table>
Perspectives
In progress...

- Validation with Google

- **GCP Studio**, a dedicated model-driven environment for GCP

- Generation of **GCP artifacts** such as JSON files and CURL scripts

- Model-driven management of GCP systems
Thank you!

stephanie.challita@inria.fr
www.occiware.org
https://github.com/occiware/GCP-Model