



Projektskizze Semesterprojekt CAS Cloud Computing

Titel: Geodatenprozessierung mit (Spot) Budget Instanzen auf Amazon EKS

1. Umfeld / Ausgangslage

Die swisstopo betreibt das Geodatenportal des Bundes. Der bekannteste Teil dieses Portals ist die Internetkarte <https://map.geo.admin.ch> (der Viewer). Dieser Viewer erfreut sich grosser Beliebtheit und beinhaltet ca. 800 Themen wie Wanderwege, Solarkataster oder Luftfahrthindernisse. Unser Team (IGEB-B) publiziert diese Daten; deren Nachführungszyklus wie auch der Aufwand zur Aufbereitung fürs Web unterschiedlich sind. Einige Daten werden manuell aufbereitet, andere stündlich automatisch nachgeführt.

Nebst der Publikation der Daten sind wir für den Betrieb und der Weiterentwicklung der Webdienste und des Viewers verantwortlich.

Der ganze Technologie Stack wurde schon länger nicht mehr grundlegend erneuert. Zurzeit wird die gesamte Architektur analysiert und überarbeitet, um eine gestaffelte Migration auf eine neue Lösung zu ermöglichen.

Einige Rahmenbedingungen dieser zukünftigen Architektur sind bereits klar: Das Geoportal des Bundes wird weiterhin in der AWS Cloud betrieben werden, die Migration wird vor allem über Microservices gestaffelt erfolgen, diese Services werden als Docker Container laufen, Amazon Elastic Kubernetes Service wird die Orchestrierung der Container übernehmen; und für Continuous Integration wird AWS Codebuild/Pipeline zum Einsatz kommen.

2. Problemstellung

Es ist schon lange ein Anliegen, dass die Publikationsprozesse von Geodaten optimiert werden sollen. Wann immer möglich, soll der Automatisierungsgrad erhöht werden.

Besonders aufwändig erweist sich zurzeit die Publikation von 3D Daten. Die manuelle Publikation der 3D Daten benötigt eigene Tools, die auf einem performanten und somit teuren Rechner laufen müssen. Ausserdem erfordert die Bereitstellung einen hohen Koordinationsaufwand zwischen der Infrastruktur und der Entwicklung. Häufig passiert es, dass Mängel in den 3D Daten erst nach beendeter Webpublikation bemerkt werden; und der ganze Publikationsprozess muss wieder von vorne gestartet werden.

Auch dem Hersteller der 3D Daten (dem Bereich Topografie) wäre es ein Anliegen, wenn er diese Daten selbst automatisch publizieren und prüfen könnte.

3. Lösungsansatz

Weil unsere zukünftigen Services auf Amazon Elastic Kubernetes Service laufen werden, finde ich es interessant, dort ein Job Sheduling aufzusetzen, dass die 3D Daten (Buildings und Vegetation) automatisch publiziert. Auf diese Weise kann ich mich im Rahmen dieser Arbeit auch gleich mit der neuen Umgebung und deren Möglichkeiten vertraut machen.

Mit diesem Piloten will ich auch gleich untersuchen, ob für die Prozessierung der Daten freie Budget Instanzen (EC2 Spot) verwendet werden könnten.

Als allgemeiner Ausblick könnte diese Arbeit helfen zu ermitteln, welche EKS Tools den Betrieb eines solchen Job Sheduling für Geodatenprozessierung unterstützen könnten.



4. Personen

Studierender	1: swisstopo, Reber, Tobias, 058 / 469 03 82, tobias.reber@swisstopo.ch
Studierender	2:<Firma, Name, Vorname, Telefon, email>
Studierender	3:<Firma, Name, Vorname, Telefon, email>
Ansprechpartner Betreuer in der Firma:	/swisstopo, Böcklin, Christoph, 058 / 469 15 14, <u>christoph.boecklin@swisstopo.ch</u>