

Verteilte Prozessierung von ALS-Daten mittels Cloud-Computing - Erste Ergebnisse einer Pilotstudie

Johannes OTEPKA, Gottfried MANDLBURGER, Willi KAREL und Bruno WÖHRER

1 Einleitung

Airborne Laserscanning (ALS) hat sich als Standardmethode zur Erfassung topographischer Daten etabliert. Aufgrund der steigenden Messfrequenz moderner Sensoren werden immer höhere Punktdichten erreicht, was zu einem stetigen Anstieg der Punktdatenmengen führt. Dadurch steigen die Anforderungen an die EDV-Infrastruktur – hinsichtlich Hardware und Software – zur Prozessierung der Daten. Im Falle großer Laserscanningprojekte ist die Bearbeitung auf einem Einzelplatzsystem kaum mehr in adäquatem Zeitrahmen möglich.

Eine Alternative dazu bietet eine Verteilung der Rechenleistung auf mehrere Rechner. Als Schlagwort steht gegenwärtig der Begriff „Cloud-Computing“ im Raum. Darunter versteht man allgemein das Bereitstellen von konfigurierbaren Computerressourcen über interne und/oder externe Netzwerke (i.A. das Internet). Mittlerweile bieten zahlreiche Anbieter (Amazon, Google, Microsoft, etc.) diverse Modelle von „Cloud-Computing“ an, welche Anwender dynamisch und je nach Bedarf nutzen können. Die Steuerung erfolgt zentral über proprietäre Tools oder Programmierschnittstellen (APIs). Darüber hinaus existieren auch im Open-Source-Bereich zahlreiche Tools, die eine verteilte Prozessierung unterstützen; ein Beispiel dafür ist das Open-Source-Projekt *gearman*. Das am Institut für Photogrammetrie und Fernerkundung (I.P.F.) entwickelte Programmsystem OPALS (MANDLBURGER ET AL, 2010) zur Prozessierung von ALS-Daten verfügt über eine C++- und eine Python-Schnittstelle. In der vorliegenden Pilotstudie wird evaluiert, inwieweit sich die hochstehenden Algorithmen von OPALS mit den Verteilungsmöglichkeiten von Python effizient verknüpfen lassen. Dazu wurde am I.P.F. eine „private Wolke“ mittels *gearman* eingerichtet. Die Ergebnisse dieser Studie und die ersten Erfahrungen mit Cloud-Computing hinsichtlich der Anwendung im Bereich des Airborne Laserscanning werden vorgestellt.

2 Cloud-Computing

2.1 Definition

Cloud-Computing ist derzeit das zentrale Thema für viele Bereiche der gesamten IT-Branche. Hardwarehersteller, Softwareproduzenten, Provider und Dienstleister nähern sich dem Thema naturgemäß auf ganz unterschiedliche Weise, weshalb eine knappe und klare Definition von Cloud-Computing schwierig ist. *Das National Institute for Standards and Technology* (NIST) veröffentlichte eine weitergehend anerkannte Definition (MELL & GRANCE, 2009), welche verschiedene Definitionsansätze bündelt (WIKI DE, 2010; CT, 2010). Neben fünf speziellen Charakteristika für Cloud Computing werden drei verschiedenen Servicemodelle und vier Liefermodelle angeführt.

Das Verteilen der Rechenlast auf verschiedene Server (wie beim Grid-Computing oder bei Peer-to-Peer-Netzwerken) ist nur ein Teilaspekt. Das zentrale Ziel beim Cloud-Computing ist das Auslagern von Rechenlast. Anstatt eigene Rechner-, Server- oder Softwareressourcen zu nutzen, werden die Ressourcen von Cloud-Anbietern genutzt. Ein klarer Paradigmenwechsel.

2.2. Servicemodelle

Im Allgemeinen werden drei verschiedene Servicemodelle differenziert, die auch unterschiedliche Anwendergruppen ansprechen sollen.

- *Infrastructure-as-a-Service (IaaS)*: Stellt dem Benutzer virtuelle Server zur Verfügung. Je nach Bedarf kann um beliebige Server-Instanzen erweitert oder verkleinert werden (z.B. Elastic Compute Cloud (EC2) von Amazon, (AMAZON, 2010)). Der Benutzer hat vollen Zugriff auf die virtuelle Hardware und kann auch selbst Anwendungen installieren.
- *Platform-as-a-Service (PaaS)*: Der Entwickler kann Anwendungen hochladen, wobei die Verteilung auf physische Server von der entsprechenden Plattformsoftware erledigt wird. (z.B. Windows Azure von Microsoft (MICROSOFT WINDOWS AZURE, 2010) oder App Engine von Google (GOOGLE APPENGINE, 2010)). Der Nutzer hat wenig bis keine Möglichkeit die Server zu konfigurieren
- *Software-as-a-Service (SaaS)*: Stellt dem Nutzer Software zur Verfügung (z.B. Google Docs (GOOGLE DOCS, 2010), Microsoft Skydrive Office Web (MICROSOFT SKYDRIVE, 2010), etc.)

Für die Prozessierung von ALS Daten können unterschiedliche Cloud-Computing Servicemodelle eingesetzt werden. Exemplarisch sollten 2 Szenarien skizziert werden

...

Alles weitere im Tagungsband