Karlsruhe Institute of Technology

Datensicherheit in der Cloud

Backups, Redundanz und Privatsphäre

Grundlagen: Redundanz vs. Backup

- Backup: Kopie von Daten und seperate Lagerung zur späteren Wiederherstellung im Falle von Verlust
- Redundanz: Daten leben stets aktuell repliziert an mehreren Orten
  - Lokale Redundanz: Innerhalb eines Datenzentrums
  - Geo-Redundanz: Über mehrere Datenzentren hinweg

Was machen die Großen? - Geo-Redundanz

- Automatischer Ablauf:
  Kunden wählen beim Einrichten eines virtuellen Speichergerätes die Art der Redundanz und die Speicherregion aus (und abhängig davon die Preisstruktur)
- Kunden haben nicht unbedingt Kenntnis über den genauen Speicherort der Daten, lediglich die Region ist bekannt (z.B. „us-west“) und bestimmte Garantien (z.B. Google: Datenzentren sind mindestens 100 Meilen auseinander gelegen)

Was machen die Großen? - Lokale Redundanz

- Beispiel Microsoft:
  Daten werden innerhalb des Datenzentrums 3 fach komplett repliziert
- Beispiel Google: Erasure Coding
  Daten werden ähnlich wie bei RAID aufgeteilt und dann auf verschiedenen Systemen gespeichert
  Systeme liegen in getrennten Netzwerk- und Stromversorgungsdomänen

Was machen die Großen? - Daten Backup

- Microsoft und Google: Scheinbar keine spezifischen Funktionen für Backups vorhanden
  → Nutzer müssen ihre Daten manuell sichern
  Aber: Günstigeres Speicherangebot für Sicherungen ist vorhanden
- Beispiel Amazon: Elastic Block Storage
  Ein virtuelles Blockdevice kann eingerichtet werden
  Davon lassen sich ähnlich wie bei CoW Dateisystemen Snapshots erstellen
  Diese lassen wiederherstellen und sogar teilen
  Aber Achtung: Snapshots sind auf Block-Ebene (könnten also auch zuvor gelöschte Daten enthalten)

Was machen die Großen? - Datenbank Backup

- Datenbanken können automatisch in festgelegten Zeitintervallen gesichert werden
- Nutzer können zusätzlich jederzeit manuell ein Backup erstellen

Backup in die Cloud: Grundbegriffe

- Versionierung
  Deltas speichern spart Platz und Kosten
- Verschlüsselung
  Daten sollen weder vom Cloud-Anbieter noch von eventuellen dritten Angreifern lesbar sein
- Assured Deletion
  Beim löschen von Daten lässt sich nicht überprüfen, ob der Anbieter diese auch wirklich löscht
  Entfernte Daten sollen jedoch garantiert gelöscht sein
  → Lösung in Kombination mit Verschlüsselung:
  Schlüssel zu jeder Datei/Version lokal halten und löschen

Backup in die Cloud: Assured Deletion & Versionierung

- Problem: Beide Konzepte schließen sich scheinbar gegenseitig aus:
  Zuerst Verschlüsseln macht Deltas unmöglich
  Zuerst Deltas macht beim Löschen einer alten Version die neuen unbrauchbar
- Lösung:
  Dateizugriff erfolgt nach Policies, jede davon hat einen eigenen Schlüssel
  Schlüssel für Dateien sind mit denen der Policies entschlüsselbar
  Löscht ein Nutzer eine Datei, wird lediglich der Policyschlüssel gelöscht

Backup in die Cloud: Data Partitioning

- Daten (z.B. Spalten in Datenbanken) aufteilen und auf verschiedenen Storage Providern ablegen
  → Kein einzelner Anbieter hat Zugriff
  → Angreifer, die Zugriff auf Daten eines einzelnen Anbieters haben, bekommen ebenfalls keinen Vollzugriff

Zusammenfassung

- Provider bieten von Hause aus viel Funktionalität um Daten redundant zu speichern und garantieren hohe Verfügbarkeit und Sicherheit (>99,99%)
- Backups innerhalb der Cloud müssen meist manuell vorgenommen werden
- Um für gesicherte Daten Vertraulichkeit zu gewährleisten müssen Nutzer selbst aktiv werden und spezielle Software oder Methoden benutzen