

symCloud

Entwicklung eines verteilten Speicherkonzeptes als Grundlage für eine Filehostingplattform

Zusammenfassung: Filehostingplattformen sind in der heutigen Zeit allgegenwärtig. Ohne einen Zugang zu einem der allgemein verfügbaren Dienste, ist heutzutage eine Zusammenarbeit in einer Gruppe von Menschen, kaum möglich. Einige Menschen jedoch haben Bedenken, ihre Daten einem Betreiber anzuvertrauen, den sie nicht kontrollieren können. Die Angst vor dem Kontrollverlust ermöglicht quelloffenen Lösungen den Einstieg in diesen Markt. Die vorliegende Arbeit beschäftigt sich mit der Konzeption einer Speicherlösung für eine derartige Software.



Ziel

Das Ziel der Arbeit ist die Entwicklung eines Speicherkonzeptes für eine neue Filehostingplattform. Diese Plattform nennt sich symCloud und ist eine neue Software, die Ideen aus verschiedenen Applikationen und Technologien kombiniert, um eine optimale Lösung für den Anwender zu schaffen.

Ergebnisse

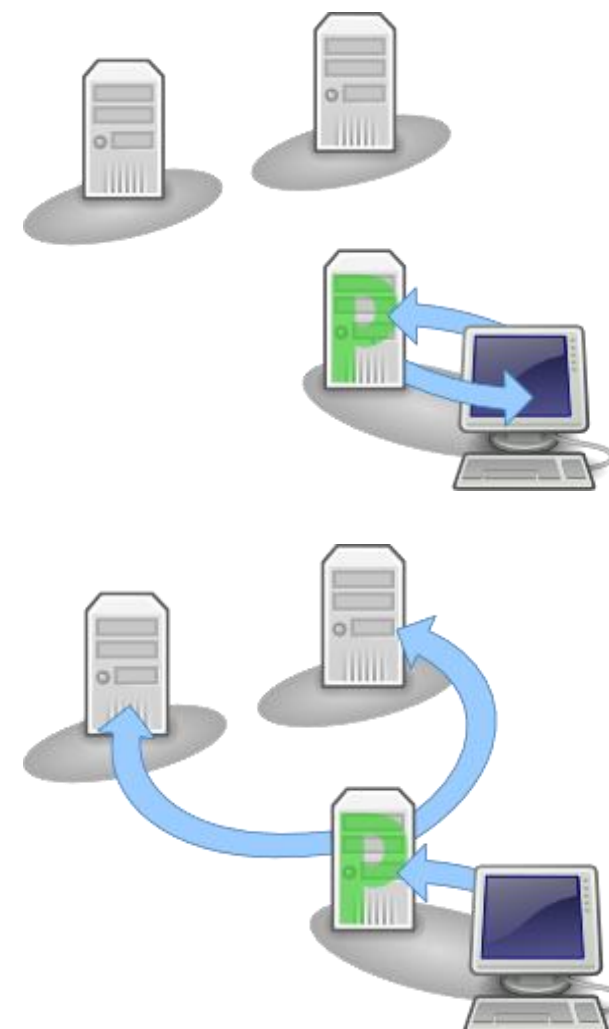
Das Ergebnis dieser Arbeit ist ein Konzept für eine verteilte Datenhaltung. Diese Speicherlösung ist unabhängig von der Anwendung, in der sie integriert ist. In einem Prototypen wurden die wichtigsten Aspekte implementiert.

Einer dieser Aspekte ist das Protokoll, das die Daten in einem konfigurierbaren Netzwerk verteilt. Es ist eine Abwandlung des von XtreamFS [XtreamFS 2014] verwendeten primärbasierten Protokoll. Es wird verwendet, um die Datensätze zu replizieren.

Stand der Technik

Für die Entwicklung des Konzeptes wurden verschiedenste Technologien bzw. Anwendungen aus folgenden Bereichen analysiert:

- Cloud-Datenhaltung
- Verteilte Daten - Beispiel Diaspora
- Verteilte Datenmodelle - Beispiel GIT
- Objekt - Speicherdienste
- Verteilte Dateisysteme
- Datenbankgestützte Dateiverwaltung



Der Client kontaktiert seinen Server und tauscht Daten mit ihm aus. Bei einem lesenden Zugriff wird kein Kontakt zu anderen Servern aufgenommen, solange die Daten in der lokalen Datenbank vorhanden sind.

Erzeugt der Client neue Daten werden diese automatisch auf die verbundenen Server verteilt. Diese Verteilung erfolgt aufgrund eines Auswahlerfahrens, mit dem die Backupserver ermittelt werden.

Weiterführende Themen

- **Performance des Replika Protokolls:** Durch verschiedene Mechanismen könnte die Performance verbessert werden.
- **Rsync - Algorithmus:** Durch die ausschließliche Übertragung der geänderten Dateiteile, könnte eben diese beschleunigt und der Speicherplatz effizienter verwendet werden.
- **Konfliktbehandlung:** Sowohl die Erkennung von Konflikten als auch die sinnvolle Lösung des Konflikts ist eine wichtige Aufgabe.
- **Protokolle:** Durch die Protokolle "Webfinger" und "PubSubHubbub" würde zum einen die Interoperabilität und zum anderen die Zuverlässigkeit gesteigert werden.



diaspora*

Literatur

XtreamFS [2014]: Under the Hood: File Replication. Online im Internet: http://xtreamfs.org/how_replication_works.php [Zugriff am: 20.05.2015].