

P2P and Grid Computing

- Exercise 3 -

Michael Scholz (Matr. 1576630)
Ulf Gebhardt (Matr. 1574373)

H 3.1:

a)

Folgende Anforderungen müssen erfüllt sein:

- (lange) Verfügbarkeit der Ressourcen
- Nutzer bekommen die Files, die sie suchen und keine Fake Files (Integrität)
- Nutzer erreichen hohe Downloadgeschwindigkeit
- System soll mit „flashcrowds“ umgehen können.

b)

Die einzelnen Bestandteile des Bittorrent-Protokolls lassen sich wie folgt aufteilen:

- *Downloader*: uploaded eigene Files (bzw. chunks) und downloaded neue Files (bzw. chunks)
- *Seed*: uploaded eigene Files (bzw. chunks)
- *Tracker*: besitzt globales Register über alle Downloader und Seeds eines Files. Es stellt somit den Locationservice dar.

Weiter wird im vorliegenden Paper über den name service (hier: supernova) gesprochen. Dieser Service stellt die einzelnen .torrent-Dateien bereit. Zusätzlich werden bei supernova Moderatoren eingesetzt, welche unter anderem die Aufgabe haben neue Dateien zu inspizieren und gegebenenfalls zu löschen. Ein „normaler“ Nutzer wird zunächst als „*moderated submitter*“ behandelt. Hierbei werden die hochgeladenen Inhalte von den Moderatoren kontrolliert. Läßt ein Nutzer regelmäßig „gute“ bzw. fehlerfreie Inhalte hoch, so erhält er den Rang „*unmoderated submitter*“. Bei diesem Rang werden die neuen Inhalte nicht mehr von einem Moderator geprüft.

c)

Eine neue Datei wird zur Verfügung gestellt, indem zunächst eine .torrent-Datei erstellt wird. Diese enthält einige Metadaten. Unter anderem zeigt sie auf den tracker. Zudem muss ein Seed die neue Datei bereitstellen, damit andere Nutzer diese beziehen können.

d)

Availability: Verfügbarkeit der Files. Meistens sind die Websites down, welche die .torrent-Dateien bereitstellen. Zudem verlassen die Nutzer meistens das Netzwerk sobald sie die gewünschte Datei heruntergeladen haben. Somit verschlechtert sich die Verfügbarkeit, da die Nutzer die Datei nicht weiter uploaden. Dies lässt sich am Beispiel des Computerspiels „*Beyond Good and Evil*“ aufzeigen. Nur 17% der Nutzer seeden die Datei über eine Stunde. Diese Prozentzahl nimmt in den folgenden Stunden weiter rapide ab. Nur 3% seeden mindestens 10 Stunden die Datei.

Data integration/pollution: Die Datenintegrität (von Inhalt und Metadaten) ist bei Supernova durch das Konzept der Moderatoren gewährleistet.

Flashcrowds: Bezeichnet die plötzliche Popularität eines einzelnen (neuen) Files. Beobachtet wurde hier ein Torrent „*Lord of the Ring III*“. Das Protokoll kam mit dem hohen Interesse an dieser Ressource sehr gut zurecht.

Download performance: Im Paper wird die Downloadperformance in die folgenden beiden Punkte aufgeteilt:

1. Effizienz: Hierbei handelt es sich um die reine Downloadgeschwindigkeit
2. Effektivität: Diese beschreibt die Anzahl an verfügbaren Dateien

Die Effizienz und Effektivität stehen hierbei in einem direkten Zusammenhang.

e)

Im vorliegenden Paper wird kein Problem mit „falshcrowds“ beschrieben. Bittorrent kann mit den „falshcrowds“ gut umgehen. Dies liegt am „*tit-for-tat-Prinzip*“. Sobald ein Nutzer die ersten chunks besitzt, beginnt er selber den upload. Somit wird die Last schnell auf viele Nutzer aufgeteilt.

f)

Zur Analyse von P2P-Systemen eignen sich keine Algorithmen. Dies liegt an der starken Dynamik, welche P2P-Systeme besitzen. Eine gute Simulation dieser Dynamik ist unmöglich.

g)

Ein zentraler Datenindex hat den Vorteil, dass die Suche einfach und effizient ist. Jeder stellt er auch einen „Single Point of Failure“ dar. Die Möglichkeit eines dezentralisierten Datenindex hingegen bietet eine gute Ausfallsicherheit. Jedoch ist hier die Suche nicht mehr effizient und somit schnell möglich.

h)

Das Problem eines zentralen Trackers kann mit einem „Distributed Tracker Protocol“ (DHT) gelöst werden. Das Protkoll nutzt UDP. Alle Informationen zu DHT findet man unter http://bittorrent.org/beps/bep_0005.html.