

Hierarchische Strukturen in Peer-to-Peer Overlaynetzen

Hannes Birck

Das Forschungsgebiet der Peer-to-Peer (P2P) Systeme wurde 1999 mit der Musiktauschbörse "Napster" auf einen Schlag bekannt. P2P Systeme sind inzwischen eine anerkannte Thematik im Bereich der Forschung und gelten zudem als heißer Trend für die Zukunft des Internet. P2P ist mehr als dezentrale Netztechnik oder File Sharing, sondern umfaßt auch Themen wie Instant Messaging, verteiltes Rechnen, Groupware, P2P Web-Dienstleistungen und kollaborative Systeme.

Das besondere an den P2P Netzen ist, dass sie als sogenannte Overlaynetze auf der Applikationsschicht angesiedelt sind. In den letzten Jahren wurde es viele neue Ansätze für P2P Overlays vorgestellt. Dabei hat die Struktur und Funktionsweise starken Einfluß auf die Effizienz der Overlays. Während Napster noch mit einer zentral organisierten Client-Server Struktur arbeitete, waren die folgenden Systeme völlig dezentral organisiert (pure P2P). Diesen *unstrukturierten* P2P Systemen haben immanente Probleme mit der Skalierbarkeit, Effizienz und Netzlast-Verteilung, was zu den Ansatz der *strukturierten* Systeme führte; diese führen ihr Routing auf Basis von verteilten Hashtabellen (Distributed Hash Tables, DHTs) durch. Diese Systeme bieten zwar die gewünschte hohe Qualität, haben aber in anderen Gebieten Nachteile; aufgrund der Dynamik (Joins & Leaves) von P2P Systemen erzeugen DHT-basierte Overlays hohe administrative Kosten zur Aufrechterhaltung ihrer komplexen Struktur.

Für die Zukunft sind die *hybriden* Systeme eine ideale Möglichkeit, um verschiedene Ansätze miteinander zu verbinden, und so ein möglichst effizientes Verhalten zu erzielen. Bekannte hybride P2P Systeme sind u.a. KaZaA und Gnutella, die beide auf dem Supernode-Konzept basieren. Besonders leistungsfähige Computer werden automatisch zu *Supernodes* erhoben um dann Server-Aufgaben (Indexing) für andere Knoten (Söhne) zu übernehmen. Es handelt sich dabei um eine hierarchische 2-Ebenen Topologie, mit den Supernodes auf der oberen und ihren Söhnen auf der darunter liegenden Ebene. Das Supernode-Konzept ist eine Verbindung von purer P2P- und Client-Server-Architektur.

Thema dieser Arbeit ist die Analyse dieser hybriden P2P Overlaynetze. Es soll dabei eine generelle Aussage über ihre Effizienz und die Vor- und Nachteile von hierarchischen Strukturen in Overlaynetzen getroffen werden. Um dieses Ziel zu erreichen, sind verschiedene hierarchische/hybride Ansätze genauer betrachtet worden. Auf Basis des KaZaA Fasttrack P2P Protokolls ist von uns ein eigenes Protokoll entworfen worden, das den Aufbau von Overlaynetzen mit beliebiger Anzahl an Hierarchien unterstützt. Mit Simulationen wurden Experimente, Vergleichsrechnungen und Analysen mit verschiedenen Strukturen und Konfigurationen durchgeführt. Beispielsweise wird gezeigt, wieviel Hierarchien optimal für ein P2P System ist, in Abhängigkeit von der Dynamik des Systems.