

„Leveraging BitTorrent for End Host Measurements“

Ralf Stange

Betreuer
Oliver Hohlfeld

Technische Universität Berlin
Wintersemester 2008/2009

Worum geht es?

- Messungen von „End“-Systemen

- und wie BitTorrent dabei helfen kann...

Messen von End-Hosts

- gewachsenes Internet
- immer mehr Dienste benötigen Informationen über Topologie und Bandbreite von End-Hosts für Optimierungen
- benötigen Messmethoden um Informationen zu sammeln
- viele Messmethoden benötigen eine Kommunikation mit dem Client
- Problem
 - Firewalls, NAT, Intrusion Detection Systeme
 - diese Kommunikation wird unterbunden

Messen mit Hilfe von BitTorrent

- sehr verbreitetes Peer-to-Peer Netz mit Millionen von Teilnehmern
- Clients erwarten eine Kommunikation von fremden Teilnehmern *und* lassen sie zu
- **BitProbes** – ein angepasster BitTorrent Client zum Sammeln von Messdaten von BitTorrent-Teilnehmern

Gliederung

- 1 Einführung
- 2 BitTorrent
 - Übersicht
 - Verbindungsaufbau
 - Choking / Unchoking
- 3 BitProbes
- 4 Ergebnisse
- 5 Zum Schluss

BitTorrent – Ein kurzer Überblick

- beliebtes „kollaboratives Filesharing-Protokoll“
- verwendet Peer-to-Peer Technik

- geeignet für die schnelle Verteilung großer Datenmengen
 - Dateien werden in mehrere **Pieces** unterteilt
 - vollständig heruntergeladene Pieces können vom Client schon wieder verteilt werden, bevor die komplette Datei heruntergeladen wurde

- BitTorrent besitzt keinen zentralen Suchdienst für Dateien

BitTorrent – Ein kurzer Überblick

Torrent bezeichnet eine Datei

Peers allgemeine Bezeichnung für Clients

Seeder Clients mit kompletten Torrents

Schwarm Sammelbegriff für die Menge der Peers, die einen bestimmten Torrent herunter- oder hochladen

Tracker Serverinstanz, verwaltet die Informationen über Peers von einem oder mehreren Torrents

Torrent-Datei spezielle Datei, enthält Hostnamen des Trackers zu einem bestimmten Torrent

BitTorrent – Verbindungsaufbau

- 1 Torrent-Datei
- 2 Client kontaktiert den Tracker und erhält von ihm eine Auswahl an Peers des Schwarms
- 3 Client kontaktiert die Peers und tauscht mit ihnen Informationen über die vorhandenen Pieces aus
- 4 dem Client wird vom kontaktierten Peer mitgeteilt, ob er Pieces herunterladen darf (`unchoke`, `choke`)
- 5 sobald ein Piece vollständig heruntergeladen wurde, wird eine `have` Nachricht an alle verbundenen Peers geschickt

BitTorrent – Choking / Unchoking

unchoke der Peer ist bereit Pieces an den anfragenden Client zu übertragen

choke der Peer überträgt keine Pieces an den anfragenden Client

- im Allgemeinen werden Clients mit schnellem Upload bevorzugt behandelt
- um neuen Clients eine Chance zu geben gibt es

optimistic unchoke

alle 30 Sekunden wird ein Client zufällig (unabhängig von anderen Kriterien) ausgewählt und auf unchoke gesetzt

Gliederung

- 1 Einführung
- 2 BitTorrent
- 3 BitProbes
 - BitProbes-Client
 - Bandbreitenmessung
 - Weitere Messmethoden
- 4 Ergebnisse
- 5 Zum Schluss

BitProbes – Ein modifizierter BitTorrent-Client

- BitProbes verwendet das unveränderte BitTorrent-Protokoll
- Peers bemerken keinen Unterschied zu normalen BitTorrent-Clients
- Ziele
 - Sammeln von IP-Adressen
 - Messungen auf Basis von TCP-Verbindungen bei unkooperativen End-Hosts ermöglichen
- Wie?
 - Ausnutzen von *optimistic unchokes*

BitProbes – Ein modifizierter BitTorrent-Client

- 1 manuelle Auswahl von beliebten Torrents
 - 2 starten der normalen Kommunikation mit den Peers
 - 3 auf Basis von *optimistic unchoke* werden von einigen Peers Pieces übertragen
- grundsätzlich werden dabei keine heruntergeladenen Daten gespeichert
 - nur Zeitverhalten (Antwortzeiten, Downloadgeschwindigkeit, etc.) und BitTorrent Protokollnachrichten werden protokolliert

BitProbes – Anpassungen

- Begrenzung des Downloads auf max. 2MB pro Peer
- Erhöhung der Zahl der gleichzeitigen Verbindungen auf 1000 pro Schwarm (typische Werte von BitTorrent-Clients liegen bei 50-100)
- Verwendung eines *Shadowtrackers*
 - zur Performancesteigerung paralleler Einsatz mehrerer BitProbes Instanzen
 - alle Instanzen leiten Informationen an einen zentralen Shadowtracker weiter
 - Instanzen synchronisieren so Peer-Adressen untereinander

BitProbes – Bandbreitenmessung

- Upload-Bandbreite von Peers
 - ausreichende Datenmenge auf Basis von optimistic unchokes
 - Verwendung spezieller Tools zur Messung (hier *MultiQ*)
 - korrekte Messergebnisse sind darauf angewiesen, dass die Bandbreitenbegrenzung beim End-Host liegt
- Download-Bandbreite von Peers
 - keine ausreichende Datenmenge für direkte Messung
 - Peers versenden `have` Nachrichten bei vollständigen Pieces
 - Schätzung einer Bandbreiten-Untergrenze auf Basis von Größe der Pieces und Abstand der Nachrichten
 - zu viele unbekannte Einflüsse für Obergrenze

BitProbes – Weitere mögliche Messmethoden

- Statistische Informationen zu BitTorrent
 - auf Basis der BitTorrent-Protokollnachrichten
 - Schwarmgrößen, Teilnahmezeiten, etc.
- Messungen durch Manipulation der TCP-Pakete (TCP-Sidecar)

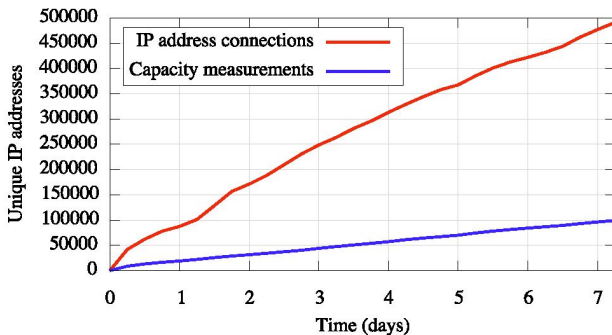
Gliederung

- 1 Einführung
- 2 BitTorrent
- 3 BitProbes
- 4 Ergebnisse
 - Praktischer Aufbau
 - Effektivität von BitProbes
 - End-Host Bandbreiten
- 5 Zum Schluss

BitProbes – Prototyp im praktischen Einsatz

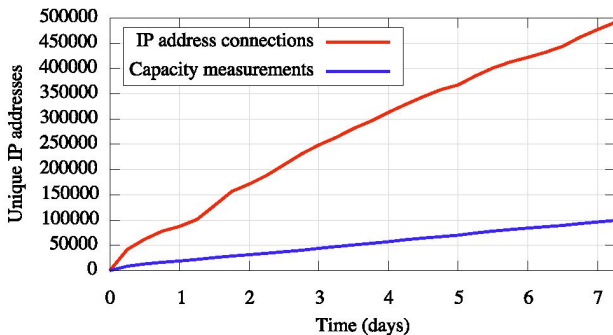
- 8 Server der University of Washington
- mit jeweils 40 Instanzen von BitProbes
- Messzeitraum vom 2.9. – 9.9.2006
- Torrents von
 - <http://thepiratebay.org>
 - <http://www.minova.org>
- Auswahl erfolgt automatisch alle 12h auf Basis der Schwarmgröße

Effektivität von BitProbes – Gefundene End-Hosts



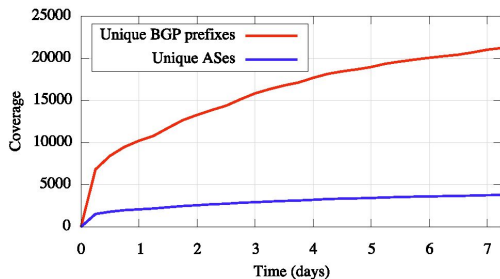
- Anzahl gesammelter **bestätigter IP-Adressen**
- Anzahl kontaktierter **Peers mit geeigneten Datendownload** (für Analyse durch MultiQ)

Effektivität von BitProbes – Gefundene End-Hosts



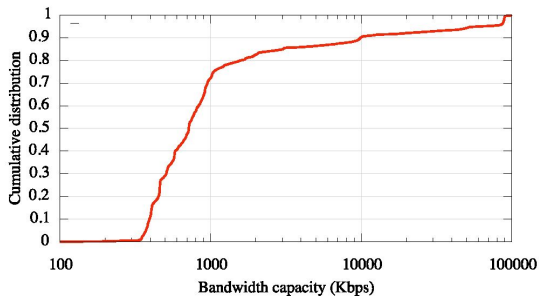
- **500.000** bestätigte IP-Adressen
- ca. 175.000 TCP-Verbindungen mit Datendownload
- ca. **96.000** zur Analyse durch MultiQ geeignet

Effektivität von BitProbes – Netzwerkabdeckung



- BGP (Border Gateway Protokoll) ist ein verbreitetes Routing-Protokoll zwischen autonomen Systemen (AS)
- Anzahl der **verschiedenen autonomen Systemen** aus denen bestätigte IP-Adressen stammen
- Verbindungen zu 3800 unterschiedlichen AS im Testzeitraum
- entspricht einer Abdeckung von über 15% bei weltweit 23.000 AS

End-Host Messergebnisse – Upload-Bandbreite der Peers



- Prozentuale Verteilung der **Upload-Bandbreite**
- bei 70% der Peers wurden Bandbreiten zwischen 0,35 und 1 Mbps gemessen
- 10% der Peers besitzen Bandbreiten von über 10Mbps
- Bandbreiten unter 0,3 Mbps sind kaum vertreten

Gliederung

- 1 Einführung
- 2 BitTorrent
- 3 BitProbes
- 4 Ergebnisse
- 5 Zum Schluss

Zusammenfassung

- Sammlung von über 500.000 bestätigte IP-Adressen von BitTorrent-Teilnehmern
- bei fast 20% dieser End-Hosts konnte eine längere TCP-Verbindung initiiert werden
- unbeeinträchtigt von Schutzsystemen (Firewall, NAT, etc.)
- keine „bewusste Kooperation“ der End-Hosts notwendig
- Bandbreitenmessung mit MultiQ als Beispiel für mögliche Auswertungen

Möglichkeiten / Probleme

- Verfahren auf andere Peer-to-Peer Netze übertragbar
- Messumfang leicht skalierbar
- Verwendbar für beliebige Messmethoden auf TCP Basis

- Belastung des Peer-to-Peer Netzes als *Free-Rider*
- Messergebnisse können stark von *nicht sichtbaren* Einflüssen verfälscht werden
- Auswahl der End-Host ist abhängig von Zielgruppe des Peer-to-Peer Netzes (Stichwort Firmennetze und BitTorrent)

- Leider keine näheren Informationen zum Prototypen BitProbes

Fragen?

Die Originalarbeit:

Tomas Isdal, Michael Piatek, Arvind Krishnamurthy,
Thomas Anderson:

Leveraging BitTorrent for End Host Measurements, 2007.

- 1 Einführung
- 2 BitTorrent
 - Übersicht
 - Verbindungsaufbau
 - Choking / Unchoking
- 3 BitProbes
 - BitProbes-Client
 - Bandbreitenmessung
 - Weitere Messmethoden
- 4 Ergebnisse
 - Praktischer Aufbau
 - Effektivität von BitProbes
 - End-Host Bandbreiten
- 5 Zum Schluss