

Content Delivery Networks and Video Delivery

*Elmar StelInberger, estelInb@elstel.org
Seminar aus Multimedia Communication*

CDN: Einführung

- alle größeren Webseiten und Dienste (Youtube, Facebook, etc.)
- verbessertes Routing, Caching & Sicherheit
- mehr als die Hälfte aller Internet-Videos verwenden CDNs
- zu Spitzenzeiten: Netflix und Youtube verbrauchen mehr als die Hälfte des Verkehrsaufkommens im Internet in Nord-Amerika

CDNs: Übersicht

- Hierarchie von Caches
- Edge Caches nahe dem Benutzer: schnellere Startzeit für Videos und eine bessere Verbindung
- DNS-System: ordnet Anfragen dem nächstgelegenen Edge-Cache zu
- Caches: Content Aware Caching, Cache-Admission & Cache-Eviction: Aufnahme- und Ersetzungsstrategie
- Algorithmen und Datenstrukturen: Bloom Filters, Leader Election, Gale-Shapely Algorithmus

Geschichte des Internet

- Beginn: ARPANET, Ziel die USA im Falle eines Atomangriffes regierbar zu halten
- erste Nutzer: Militär und Universitäten
- Verbindung einzelner Computern miteinander
- WWW: Content-Centric, Tim Berners-Lee, CERN
- Heute: Internet ist inhaltszentriert, man weiß gar nicht mehr genau wo die Inhalte, die man anfordert, verarbeitet und gespeichert werden

CDNs: Overlay Networks

- Routing Overlay
- Caching Overlay
- Security Overlay

Overlay $\triangleq$ virtuelles Netzwerk; → Modern Overlay
unterschiedliche Anforderungen: Video on
Demand, Live Video, interaktive Anwendungen
(Edge Comp.)

verbessertes Routing

Internet Backbone Routers:
gleiches Ziel → gleicher Weg

Routing Overlay: eigene Pfade; Optimierungsmögl.:
2,5x USA-Asien, 1,5x innerhalb N-Amerikas

Internetstruktur: baumartig, Backbones

- WorldCom Outage, 2002
- Cable and Wireless un-peered in 2001
- SEA-ME-WE4 Kabel in 2010 (Europa-naher Osten)

Video: Reflectors – über zwei Pfade doppelt senden

CDNs: Sicherheit

- Denial of Service Angriffe: widerstandsfähiger aufgrund der Größe: CDNs von Drittanbietern wie bei Netflix
- gebündelte Security-Expertise
- Kennzeichen von DoS-Angriffen: eine kleine Menge an Clients ist verantwortlich für den überwiegenden Großteil der Anfragen
- Origin Servers brauchen nur für das CDN zugänglich sein

DNS-Resolution

- Recursive Local LDNS (Local Domain Name Servers), üblicherweise vom ISP
- Authoritative Name Servers: powered by the CDN
- Global Load Balancing: Verbindung & Auslastung (Verzögerung, Paketverluste, Durchsatz)
- TTL – Time to Live

DNS-Resolution

- Private LDNSes: 400km übliche Entfernung
- Brasilien, Australien, Argentinien: 7000km
- Erdumfang: 40.000km

OpenDNS, Google Public DNS (1700km)

Worst Case! – Verbesserungsbedarf: >3000km

EDNS0 Client-Subnet Extension: User Mapping System

Client-IP: Query: /24-Präfix, Antwort: bspw.: /20-Präfix

RTT: $\frac{1}{2}$ reduziert (Dldzeit), Time-To-First-Byte: -30%

Video-Startzeit / EDNS0 Anf.

- 1 Sekunde Wartezeit: -7% Abonnements
- 4 Sekunden: 25% der Benutzer brechen ab
- sogar 100ms machen einen Unterschied für Einkünfte
- Suchmaschinen honorieren rasche Antwortzeiten

EDNS0-Erweiterung:

- #DNS-Queries: 870K → 1,17 Mio

Global Load Balancing

Gale Shapely Algorithmus

Mapping: Client-Anfragen – Server Cluster

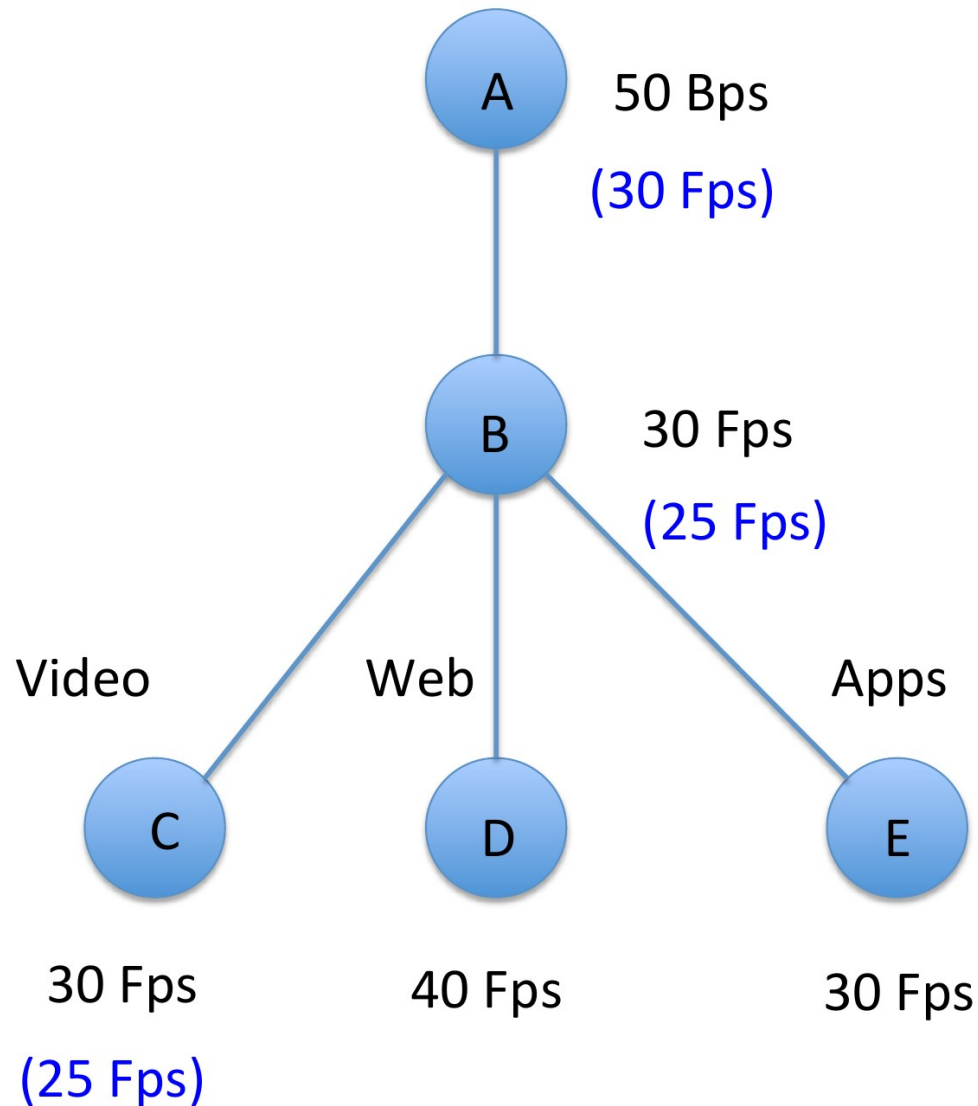
Client Anfrage: Mapping Unit (Client IP-Präfix, Traffic Class: Video, Web, Application, Download)

ursprünglich Verheiratung: Männer mit Frauen

beide Seiten: Präferenzliste, Komplexität: $O(n^2)$

Stable Allocation: für jedes alternative Paar: min. einer mit bestehendem Mapping mehr zufrieden

Ressourcenbäume



Bps – Bytes per second

Fps – „flytes“ per second
(all non-networking
resources)

Local Load Balancing

Zuweisung zu einem bestimmten Server innerhalb eines Clusters

- erster Schritt: Abbildung auf Teilmenge von Servern gemäß „Seriennummer“
- Seriennummer ~ Content Provider / Kunde
- Dateien einer Webseite zusammen speichern!
- zweiter Schritt: Hashing: für gewisses Datum wird ein bestimmter, aber immer der gleiche Server ausgewählt (häufige Dateien mehrere Server)

Local Load Balancing

DNS-Query: genaue Datei nicht bekannt, nur der Domainname!

einfachste Lösung: Anfrage und später Ergebnis an/von dem Server weiterleiten, der die entsprechende Datei tatsächlich cachet

→ Netzwerkbandbreiten innerhalb des Clusters sehr groß

Alternativen: HTTP Redirect, URL redirection (DASH-.mpd), Anycast

Cache Admission Strategy

most widely used: NHIT: long tail pollutes the cache

$\frac{3}{4}$ nur 1x nachgefragt, 90% weniger als 4x

→ Disk Load Reduction, Hit Rate, SSD: Wear-Out

Reset Period: 90% 6 hour inter-arrival time

Cache Hierarchien:

- LCE – Leave Copy Everywhere
- LCD – Leave Copy Down (oft bester Kompromiß)
- MCD – Move Copy Down

Cache Eviction Strategy

LRU outperforms LFU by 5-10% für Δ Cachegrößen

Segmented Least Recently Used (SLRU)

probationary Part / protected Part

Adaptive Replacement Cache (ARC)

SLRU mit adaptiven Segmentgrößen

Ghost List mit Hashes – gleich lang wie Segment

Verbesserungen: TTL – Time to Live für GL-Einträge

Cache Hierarchies

Origin Server – Parent Caches – Edge Caches

1GB, 10GB, 100GB, 1TB, 10TB (Youtube)

- Geo Split Caching (Baumstruktur)
- Object-based Split Caching (fester Zusammenhang)
- Hybrides Caching

Ziel: bestmögliche Antwortzeiten ($\neq$ ~ Hit Rate)
→ response latency

Content Aware Caching

music is popular over weeks, news only on same day

early video popularity unsuitable for prediction

music videos: ARC/NHIT

entertainment, people & blogs: SLRU/LCE

popularity changes over time / even diurnally

historische Segmentgrößenanpassung

Segmentgrößenanpassung aufgrund von Vorhersage:

ARIMA: Autoregressive Integrated Moving Average

ACDC – Content Aware Caching

ACDC – Adaptive Content-Aware Caching

momentane Bestandsaufnahme → Segmentanpassung

Δ Kategorien: gemeinsamer probationary Part

Eviction Strategy:

- SGL: Smallest Ghost List (GL rarely used)
- LGL: Largest Ghost List (irrelevant items on GL)
→ TTL of Ghost List (Time-to-Live)
- RSGL, RLGL: divide by size of GL

ACDC – Content Aware Caching

RLGL am besten für die meisten Kategorien

SGL: best for Sports & Comedy

Problem: kleine Segmente werden immer kleiner

eigene Verbesserungsvorschläge:

- Vergleich von Sports & Comedy mit SGL
- andere in Beziehung zu Mittelwert setzen,

Mittelwert wird ermittelt indem temporär für alle Kategorien bspw. SGL angenommen wird

Bloom Filter

approximieren nicht aufzählbare Menge:

nur ist-Element? – Überprüfung

aufgrund von Hashwerten, Bitfeld mit Bit für jeden Hashwert → False-Positive Rate

Hashing: k mal → geringere False-Positive Rate

- Cache Filtering (NHIT)
- Cache Summarization (Counting Bloom Filter)

m Bits, k Ziehungen, n Elemente

Bloom Filter

m Bits, k Ziehungen, n Elemente

$E(B[h(o)] = 0) = (1 - 1/m)^k \rightarrow p \sim \text{False Positive Rate}$

$$p = \left(1 - \left(1 - 1/m\right)^{kn}\right)^k \approx \left(1 - e^{-kn/m}\right)^k$$

$$k \approx (\ln 2)(m/n), \quad p \approx 2^{-k} \approx 2^{-\frac{m \ln 2}{n}}$$

Leader Election

avoid single point of failure:

Load Balancing, Zuweisung von Seriennummern
(Kunde/Content Provider) → Servermenge

broadcasting health values (connectivity,
sufficiently recent data, free resources)

für Timeout nichts gehört / periodisch

at-least once Semantik → mehrere Partitionen

at-most once Semantik (avoid adverse effects)

Outlook & Conclusion

Top Content Providers: majority of internet traffic
for CDNs predicted: 71% of global traffic by 2021
video streaming, web pages, social networks
basic research decades before first CDNs
Google, Facebook, Microsoft: in-house CDNs
third party CDNs: Limelight, Akamai, Level 3,
MaxCDN, Chinacache
reduce redundant traffic over the web

New Developments

Software Defined Networks – get current connectivity values

→ Routing, Global Load Balancing

ALTO Protokoll, Network ↔ Application

Edge Computing: uncacheable Content / Application Logic

Cloud Technology: rent virtual machines for any kind of period, may be hosted/migrated anywhere

PA-CDNs: P2P & CDNs

P2P – Networks: bekannt: Piratebay, Bittorrent
self scaling property: more clients → more resources
use for popular content

Probleme: Startup-Delay, QoS, Firewalls,
heterogenous Resources, incentives for user
participation

P2P-networks: also content centric

PA-CDNs: Akamai, ChinaCache, Xunlei

Fragen, Anregungen?
Literatur: siehe Artikel