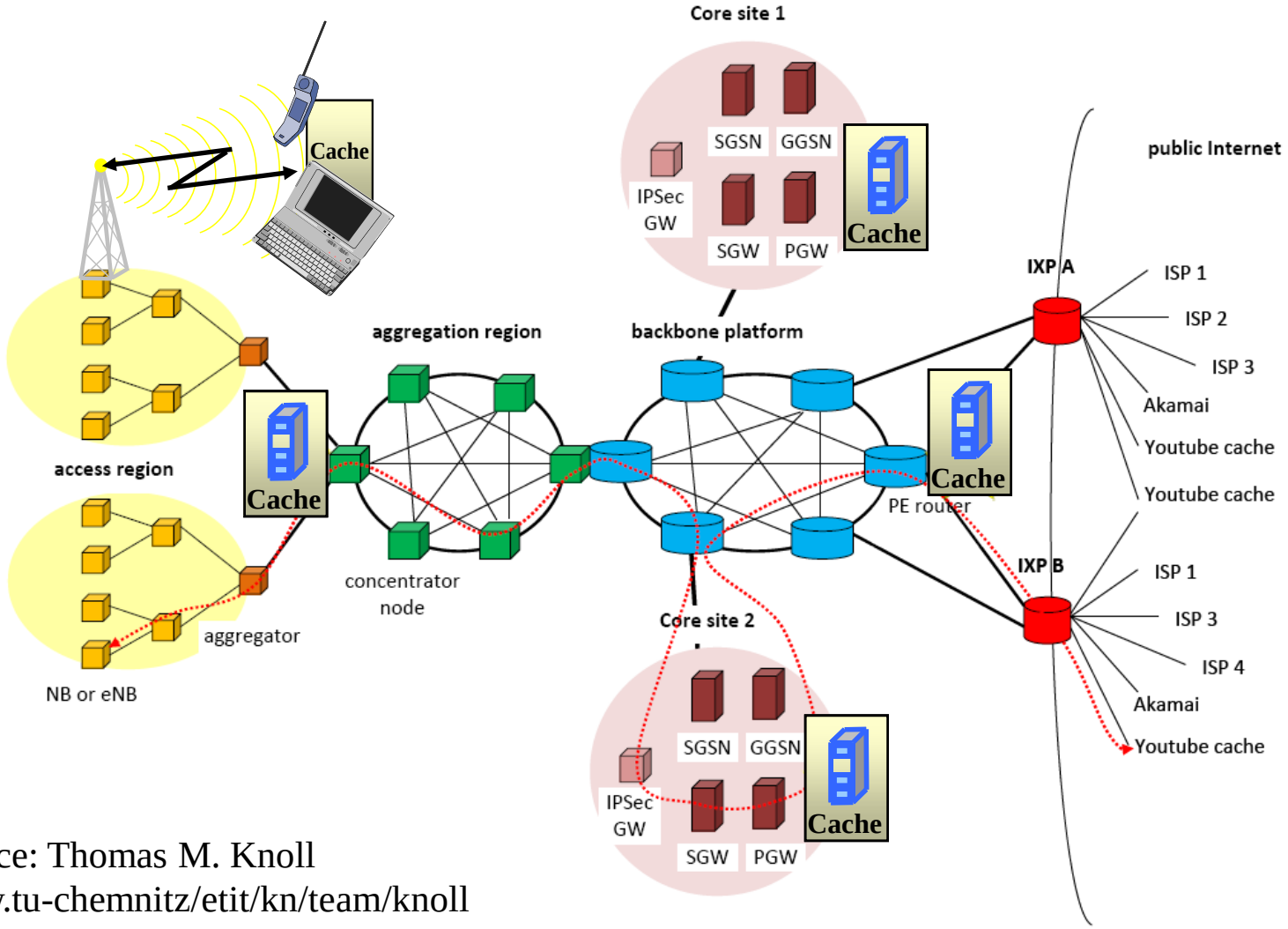
The Effect of Caches for Mobile Broadband Internet Access

Jochen Eisl, Nokia Siemens Networks, Munich, Germany
Gerhard Haßlinger, Deutsche Telekom Technik,
Fixed Mobile Engineering, Darmstadt, Germany

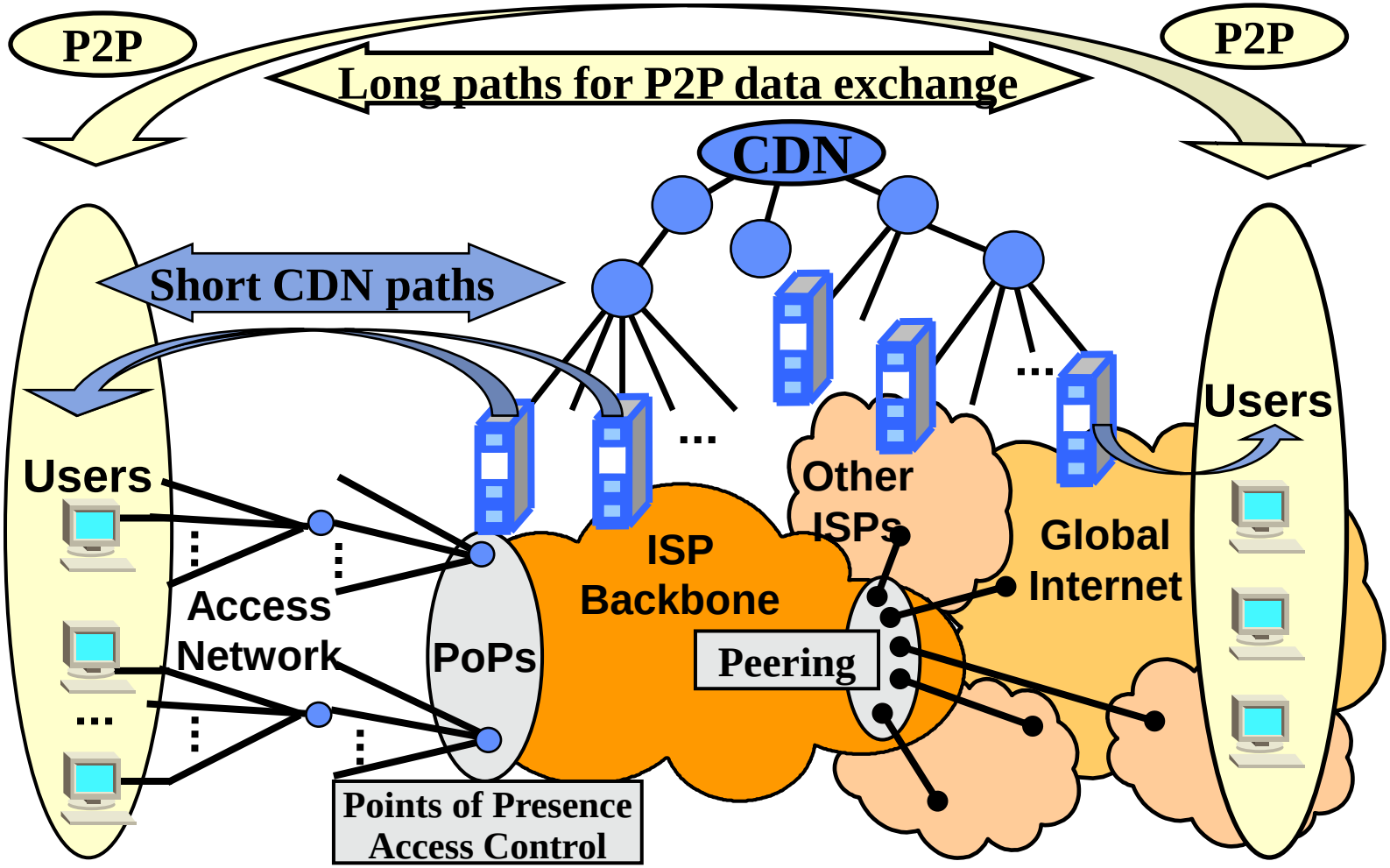
- IP-based content delivery: CDN & cache architecture
- Impact of
  – access pattern
  – application mix
  – content provider preferences & standardization
  on efficiency of caching
- Conclusions on cache efficiency in fixed/mobile networks

Overview: Cache Locations in Mobile Networks

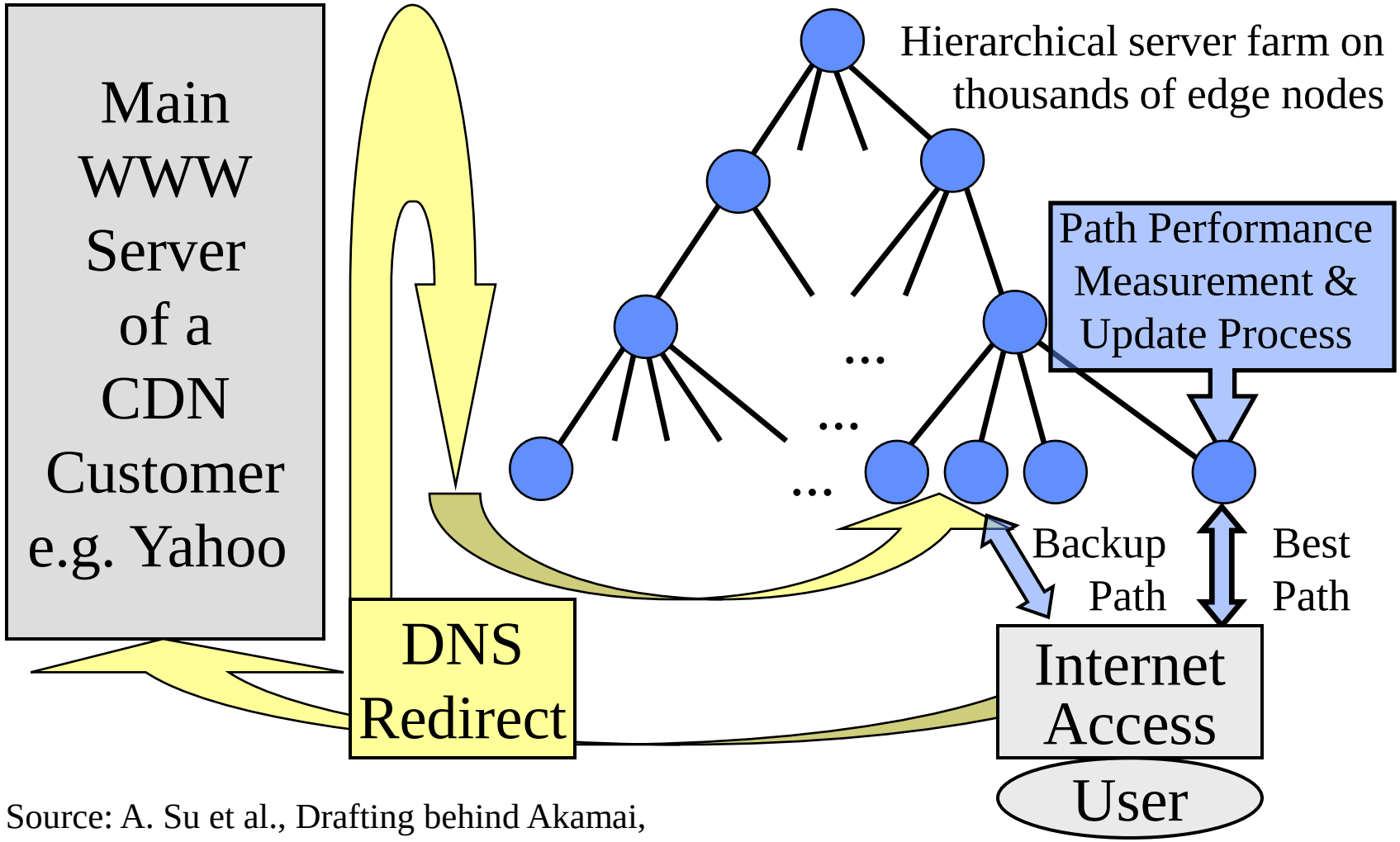


Source: Thomas M. Knoll
www.tu-chemnitz/etit/kn/team/knoll

Content Distribution: CDN ↔ Peer-to-peer overlays



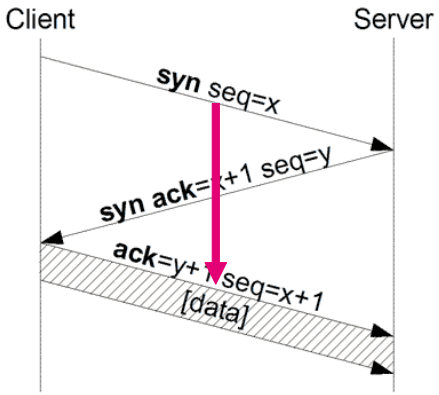
Content Delivery Overlays (CDN, Akamai, Limelight...)

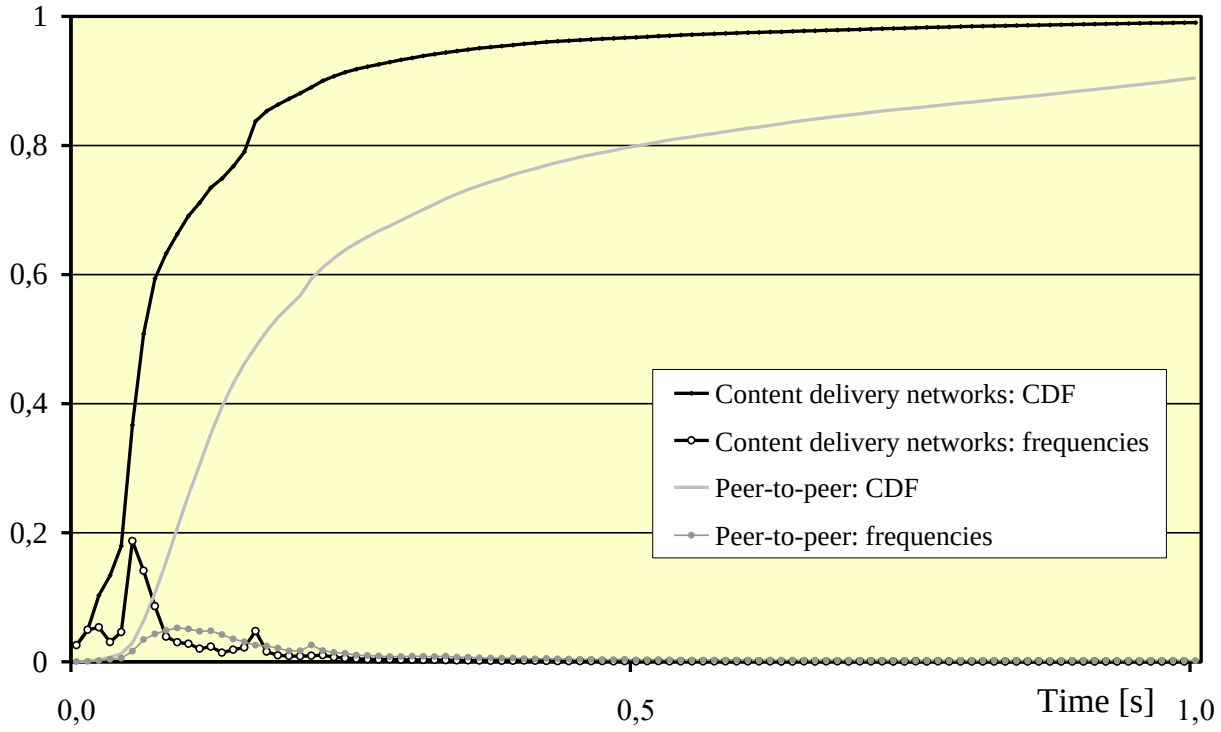


Source: A. Su et al., Drafting behind Akamai,
IEEE/ACM Trans. on Networking 2009

Link Measurement of RTT-Delays for CDN and P2P

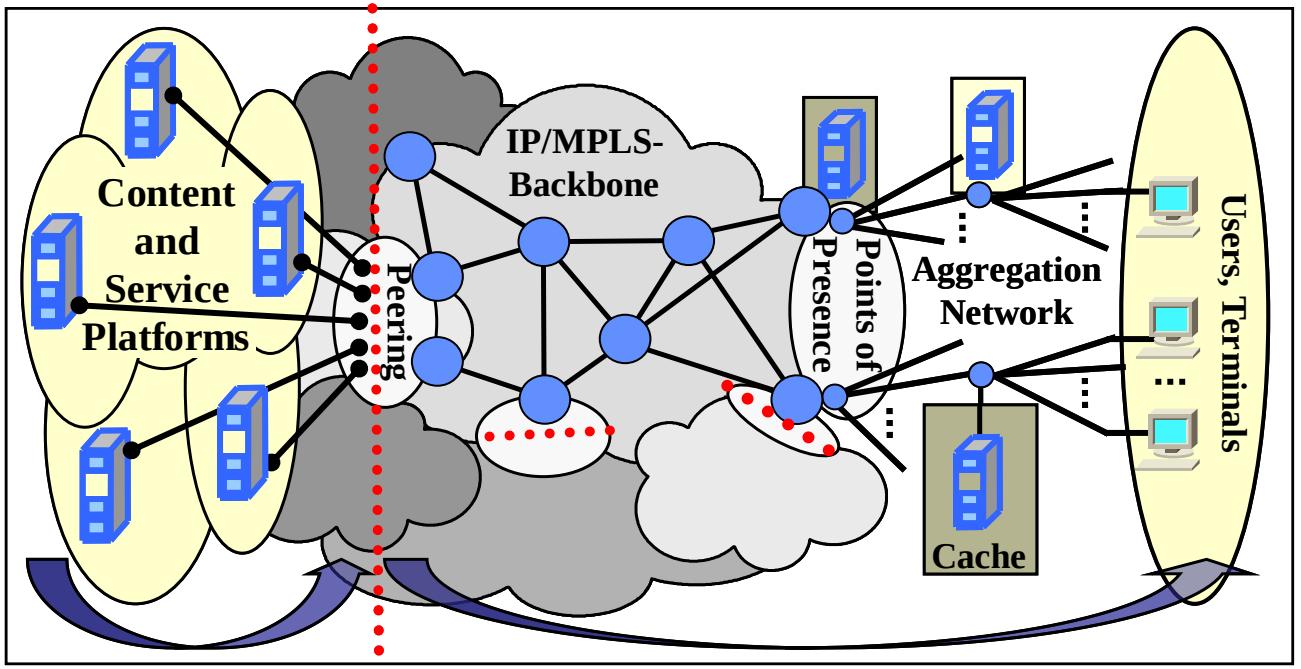
Method based on Round Trip Time (RTT) in TCP-Handshake





Cumul. Distrib. Fcts. of Round Trip Times:
Longer RTTs & paths for P2P than via CDN
Mean RTT: 0.125s for CDN; 0.33s for P2P

Overview: Caching & CDN in broadband access



Caching is applied in global content delivery networks and in network provider platforms of large ISPs …
but usually without cooperation between them!

Working group on CDN interconnection at the IETF

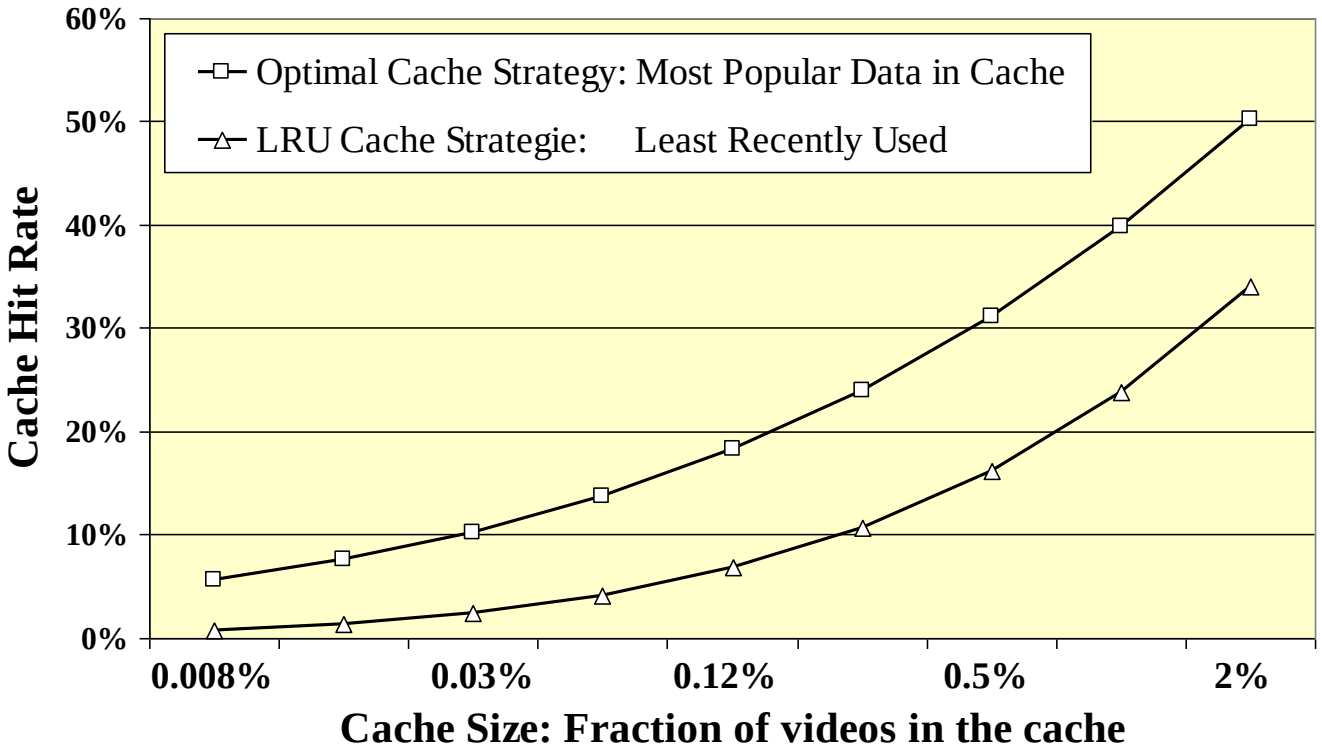
<datatracker.ietf.org/wg/cdni/charter>

Zipf Laws Determine Access to Popular Content

Zipf "80÷20" laws $A(R) = \alpha R^{-\beta}$ ($\alpha > 0$; $0 < \beta < 1$) are observed

- for access to videos on IP platforms, books, DVD sellings etc. (YouTube, America Free TV, BBC iPlayer, Amazon …)
  → L. Breslau et al. (1999): many cases where $0.64 < \beta < 0.85$;
  → P. Gill (2008); → M. Eubanks (IETF Plenary 2008)
  → Caches of small size are efficient for Zipf law accesses!
- for file sizes, transfer times → Crovella & Bestavros (1997)
- for IP connectivity (web links, BGP routes between AS nodes)
  → A. Barabási and R. Albert (1999):
  Scale-free networks → build randomly; small world effect
- for traffic volume per session and per broadband access users

Cache Efficiency e.g. for YouTube Videos



Evaluation of statistics on YouTube accesses in Dec. 2006:
3.7 billion accesses included on 1.65 million files

Source: M. Cha et al., I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system, Internet measurement conference IMC, San Diego, USA (2007)
Basically confirmed in recent YouTube evaluation at TU Munich (L. Braun et al. 2012)

Cacheability of popular Internet data

- An essential portion of IP traffic uses HTTP protocol, most of which is marked as being cacheable, often with expiry date
- Requests focus on most popular content ⇒ small caches are efficient
  Zipf law ⇔ 90÷10 rule: 90% of requests address only 10% of content
- Some content providers/CDNs support caching, e.g. software updates
- … others don't: Personalised communication with user
  ⇒ makes content identification difficult for cache manager;
  no standard feedback & control between cache ↔ content provider
- Some content providers/CDNs have business relations with content owners and/or users but often without involving network providers

Conclusions on Caching Efficiency

- Caches reduce traffic and delays ⇒ Improved QoS
- Popular websites are supported by global CDNs,
  with limited cooperation to CDNs of large network providers
- Caches on user devices can avoid ~20% of traffic
- Zipf law popularity of content ⇒ small caches are efficient
- Dynamic data, business models of content providers and
  lack of standardized cache communication as main drawbacks
- Combined CDN / P2P networks to achieve highest throughput

Conclusions on Caching in Fixed ↔ Mobile Networks

- Traffic via mobile platforms makes only a few % of fixed network IP traffic ⇒ Separated fixed / mobile caching but mobile network traffic is catching up with high growth
- Transmission capacity over air interfaces is most expensive
  ⇒ Caching on end devices should be fully exploited
  User driven prefetching strategy seems useful
- Locally relevant content in mobile networks, e.g. augmented reality applications and specific mobile content (apps, software downloads etc.)