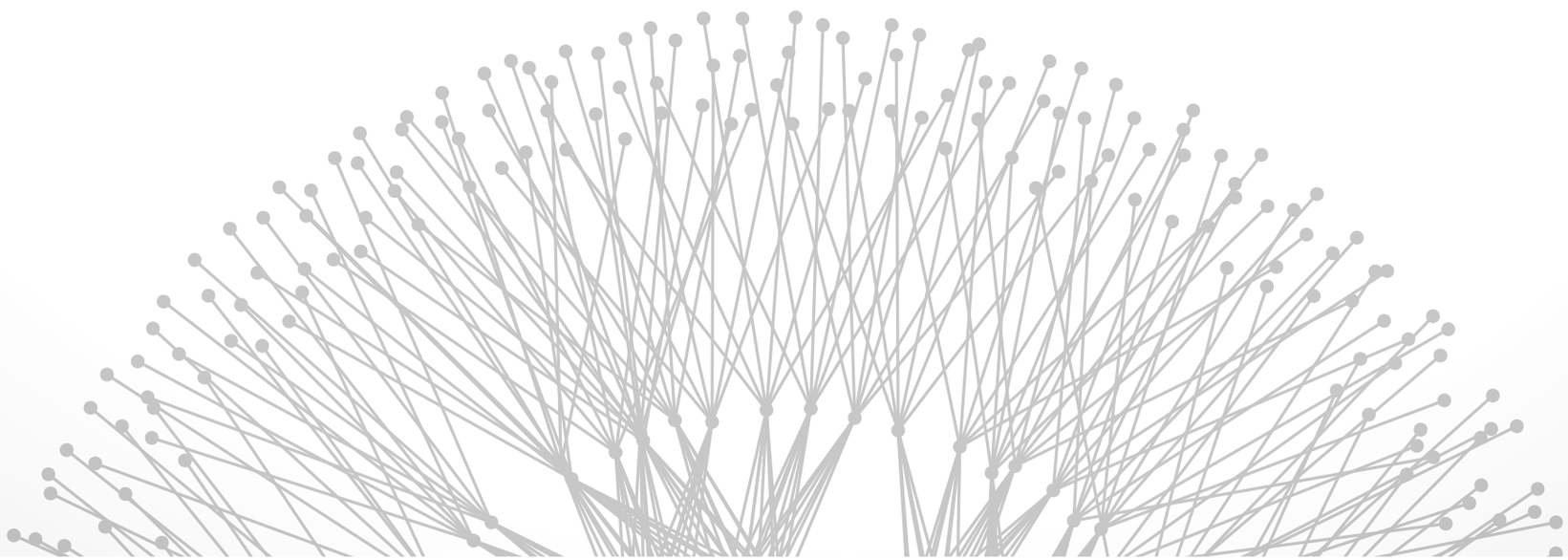
P2P, Content Distribution and Overlay Networks

Brighten Godfrey

CS 538 November 19 2013



Definitions

Peer-to-peer system: participants have the same functionality and role in the system

- ...as opposed to client-server architecture
- Commonly used to imply file sharing but also used in other contexts (e.g., "BGP peering session")
- At transport level: peer = both client and server

Overlay network: a virtual network whose links are end-to-end paths in another network

Peer-to-peer networks: Intersection of the above two

- Or, can also mean "file sharing systems"

In the beginning...

Napster (1999)

- Centralized index server to find the right peer
- Peer-to-peer file transfer

Gnutella (2000)

- Fully decentralized P2P indexing: scoped flooding
- Problems?

Freenet (1999)

- Goal: censorship-resistant key-value content store
- Routing: heuristic clustering of similar keys

In the beginning...

Napster (1999)

- Scales poorly, subject to attack (or take-down!)

Gnutella (2000)

- Flooding wastes resources, can't find all results

Freenet (1999)

- Heuristic key-based routing promising, but no guarantees

Is there a fully decentralized storage system which is guaranteed to find desired results?

Key properties of a DHT

Hashtable interface (fast *put*(*k*,*v*), *get*(*k*)=*v*)

- Freenet: get() might not find results
- DHT: guaranteed to find results, relatively quickly

Scalable

- Low memory / communication
- Uses consistent hashing: transfers in expectation 1/*n* of objects when a node leaves/joins

Resilient and decentralized

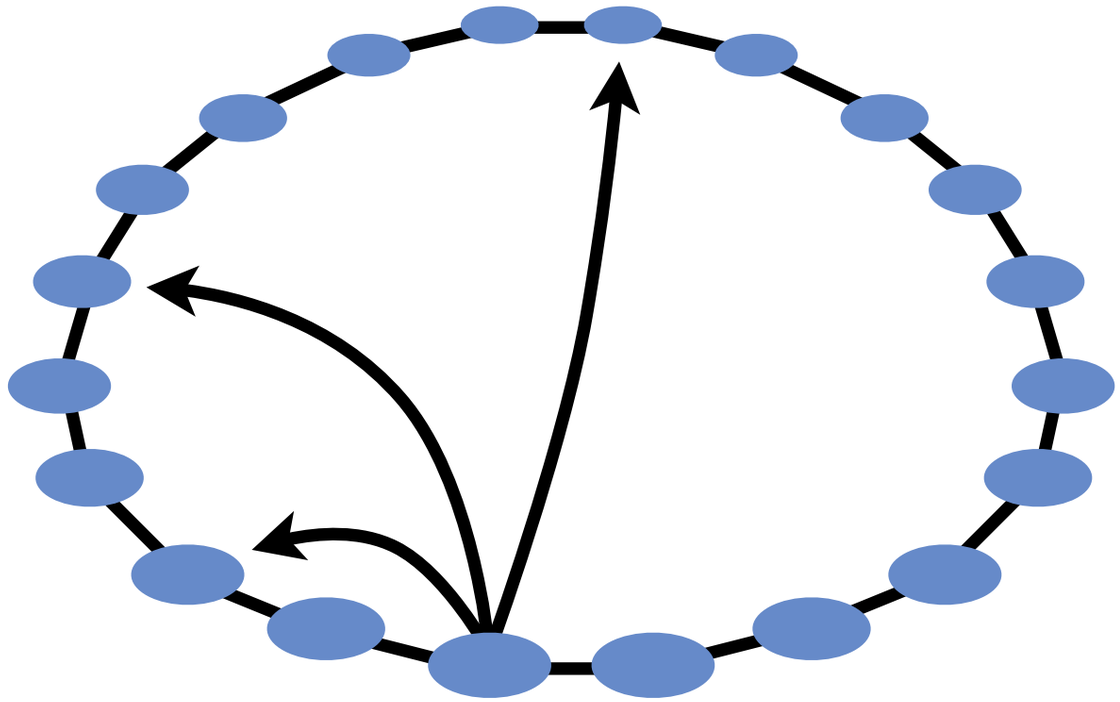
- Still works if, say, 50% of the nodes suddenly fail
- No centralized index server which could be attacked

DHTs: carefully structured

Greedy routing based on distance in keyspace

(Where did we see greedy routing before?)

- Geographic routing
- Small world models
- Grid / torus



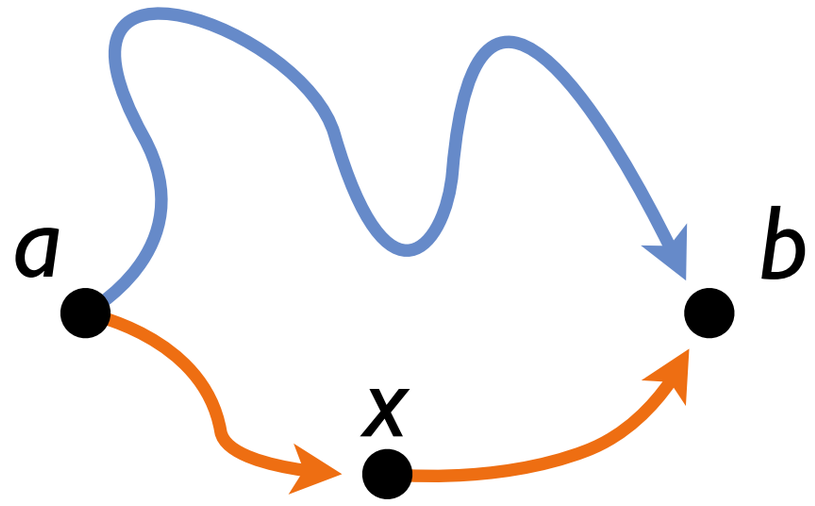
What does the DHT topology need for routing...

- ...to work?
- ...to work well?

In the other beginning...

Internet routing is suboptimal

- Observed delay $d(a,b)$ may not be best possible (why?)
- Key: Internet does not obey the triangle inequality
- i.e. it can happen that: $d(a,x) + d(x,b) < d(a,b)$



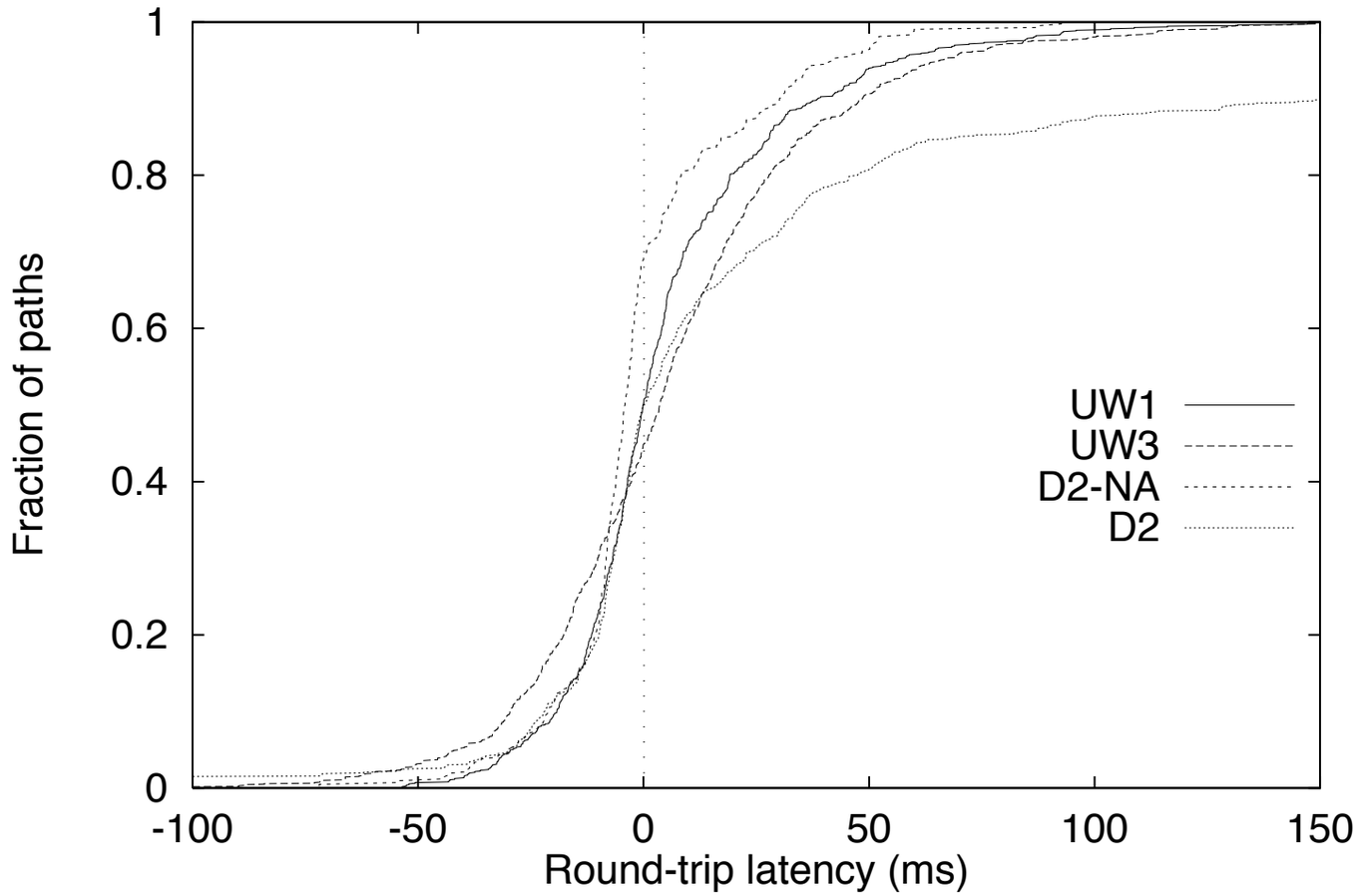
Idea: Improve it with an overlay

- Find a good point x to relay packets!

In the other beginning...

Idea: Improve it with an overlay

- "E2E effects of Internet path selection", [Savage, Collins, Hoffman, Snell, Anderson, SIGCOMM 1999]



CDF of difference between mean RTT on Internet's default path, and best mean RTT on an alternate path

- Technique used in production in Akamai's CDN

Building on the Internet's services

Common theme of many overlay networks: provide more advanced services than the Internet provides

- Much easier to deploy new functionality at hosts
- The Internet doesn't even know what's happening to it

Examples

- RON: more reliable, efficient routing
- DHT: flat name routing and key-value store
- i3: indirection, mobility, middlebox support, ...
- Content distribution: a kind of time-delayed multicast

DHT & overlay in the real world

Deployed systems

- Content distribution: Akamai, CoralCDN
- Swarming: DHT for BitTorrent distributed tracker (Vuze)
- File sharing: DHTs in Kad, Overnet/eDonkey
- Storage: Amazon Dynamo
- Botnets: Storm botnet's command & control delivered via DHT

Big impact on many research systems & papers

- Many ideas from DHT / overlay research incorporated into other work, if not entire DHT system

On to CDNs…

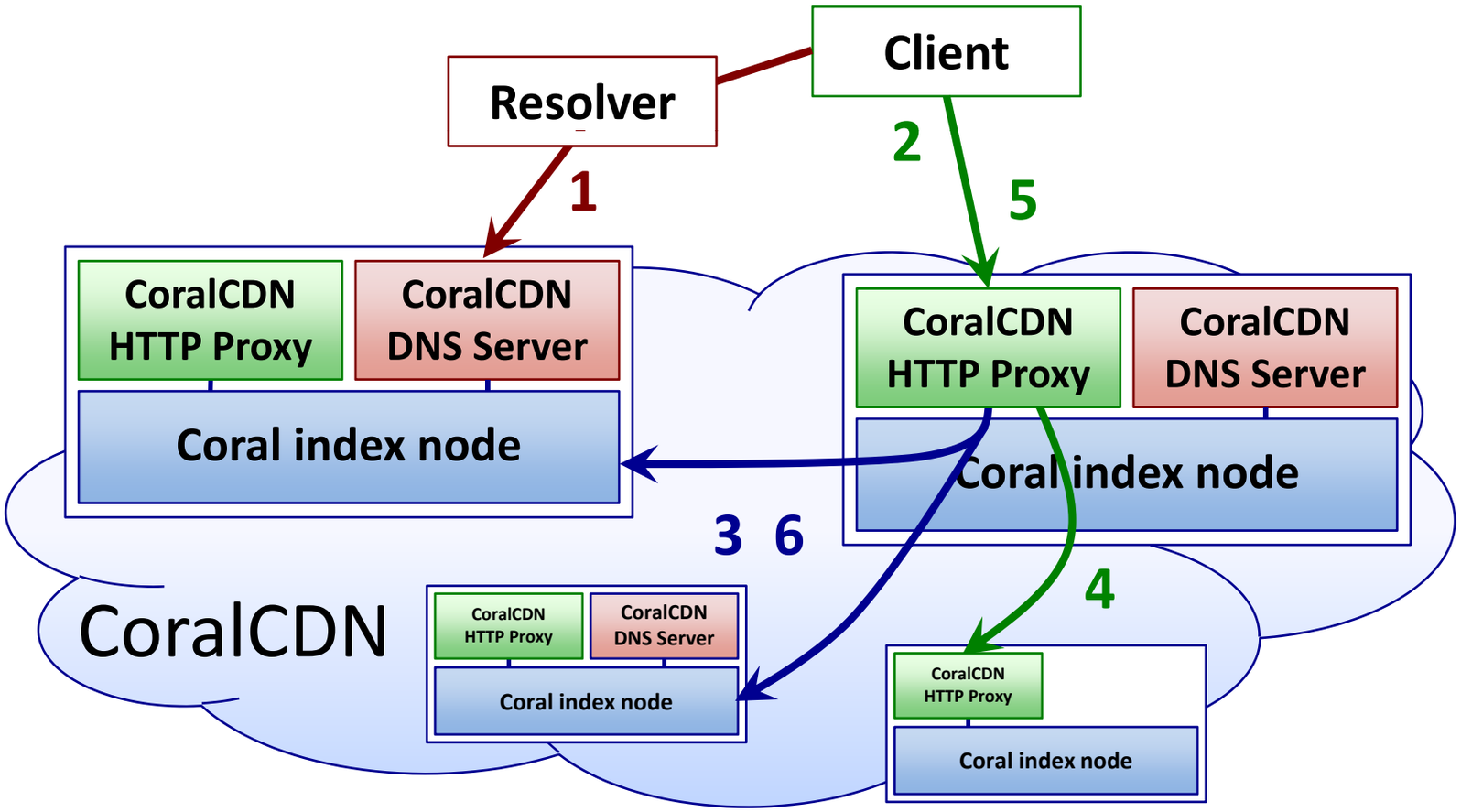


Figure from Freedman, NSDI'10

A Case for a Coordinated Internet Video Control Plane

Liu, Dobrian, Milner, Jiang, Sekar, Stoica, Zhang

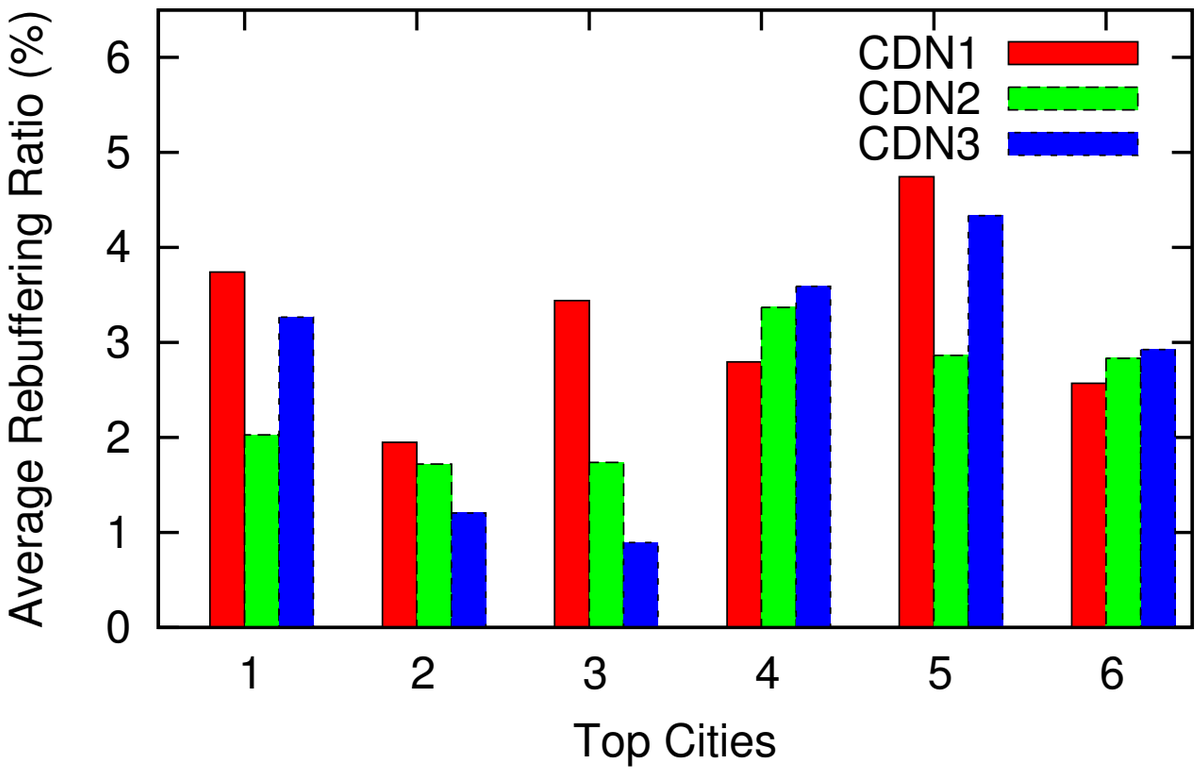
SIGCOMM'12

Key take-away points

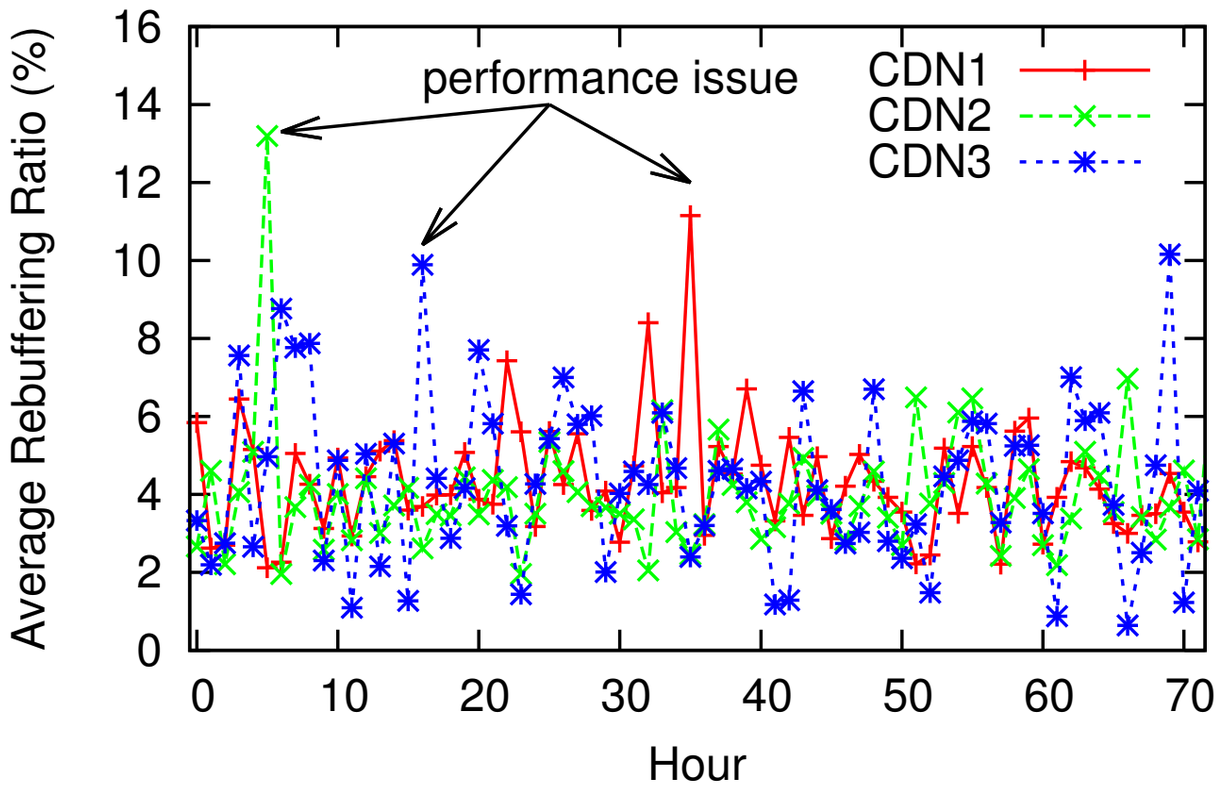
Key take-away points

Significant variability across attributes

- CDNs (even within the same city), locations, delivery rate, and time



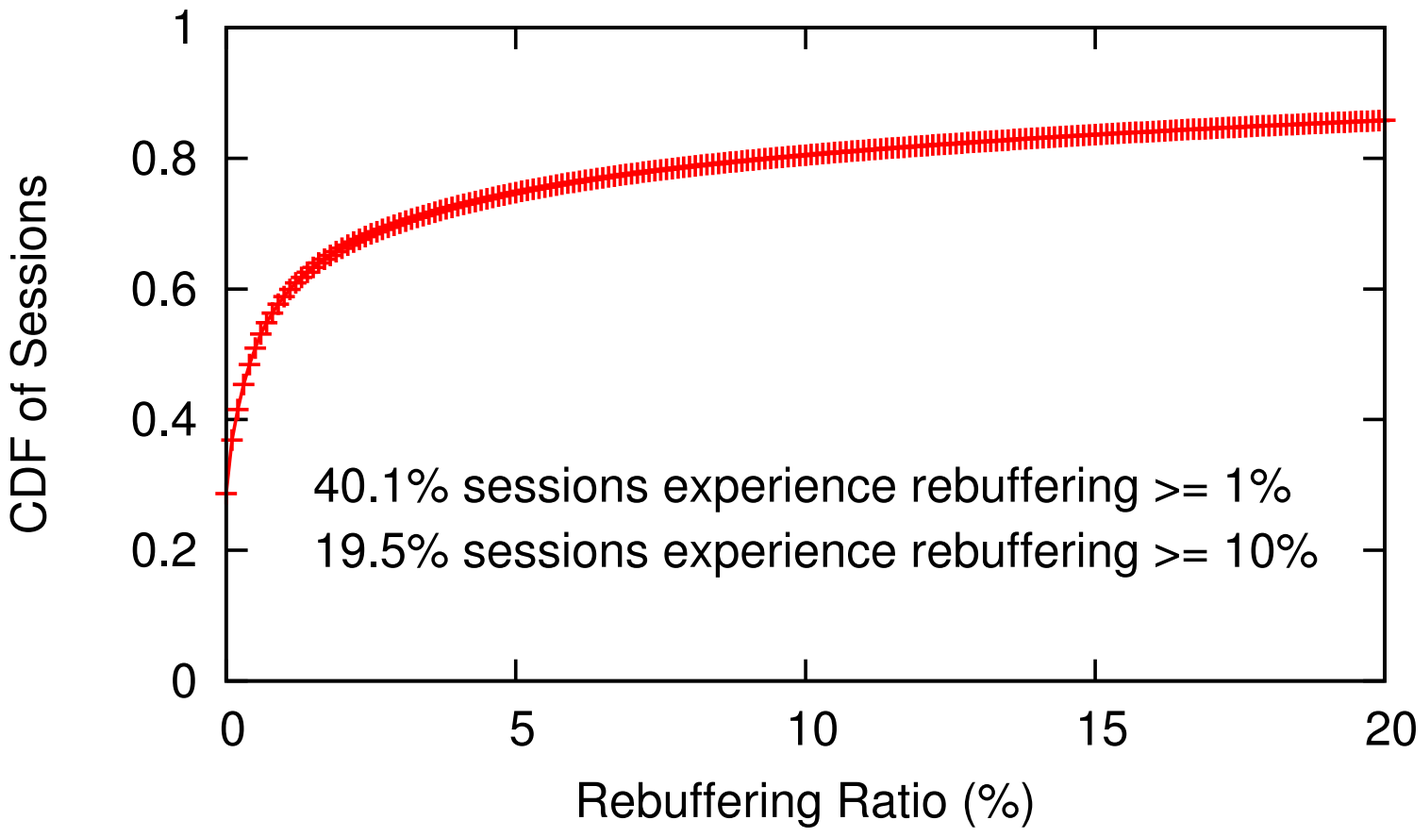
(a) Rebuffering Ratio



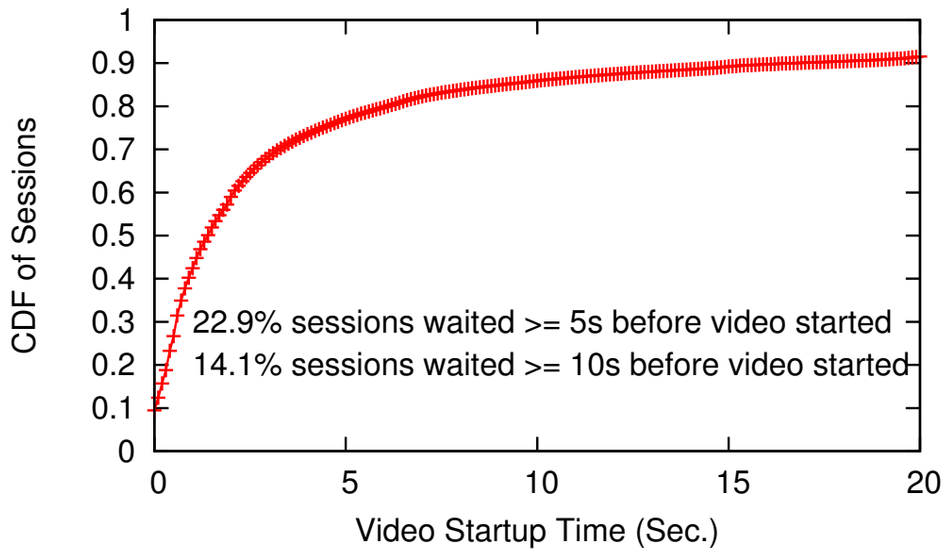
(a) Rebuffering Ratio

Key take-away points

Significant fraction of users with poor experience



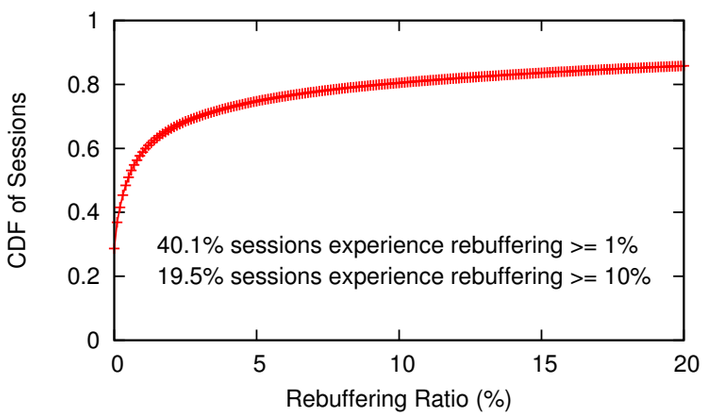
(a) Rebuffering Ratio



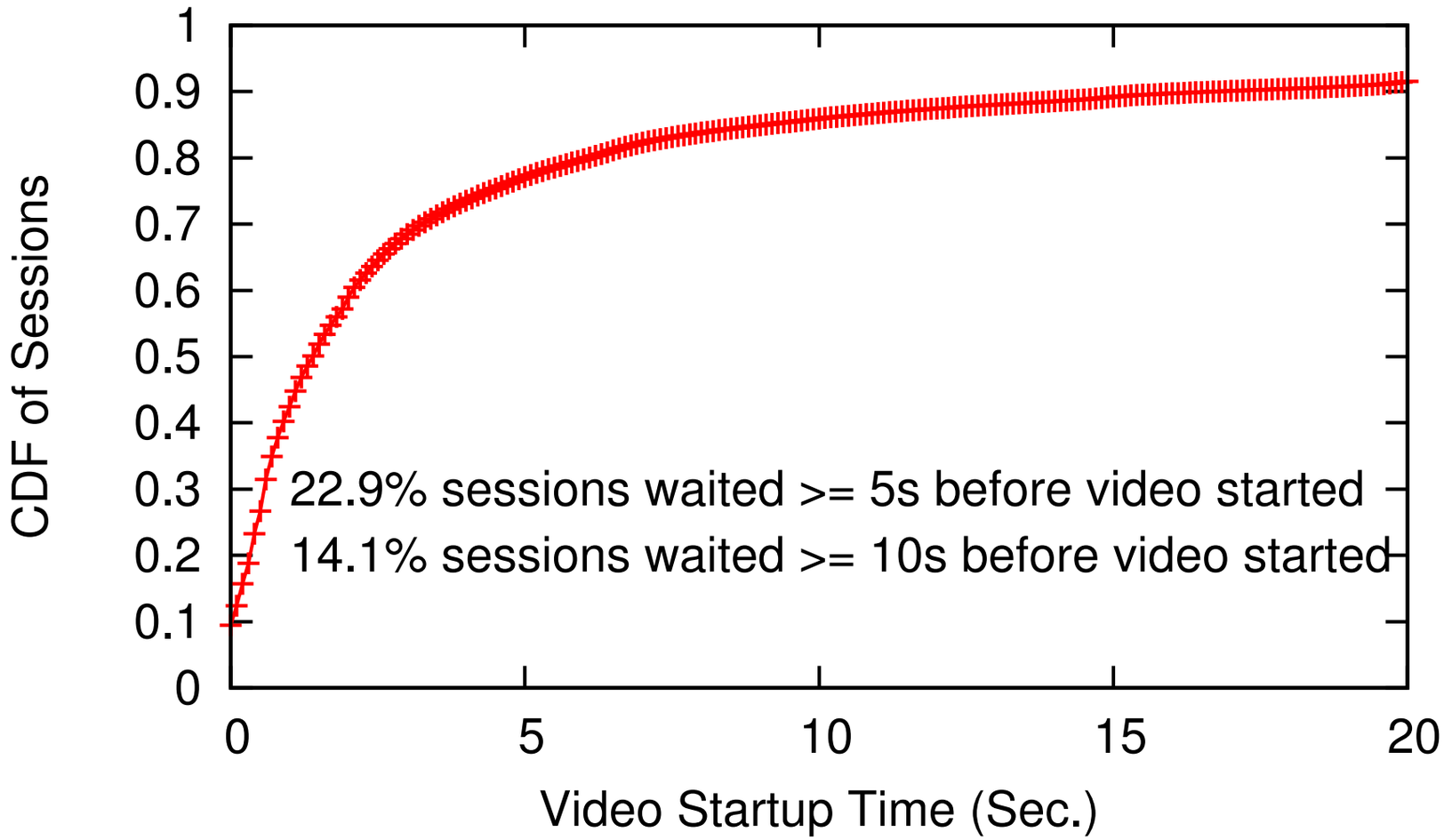
(b) Video Startup Time

Key take-away points

Significant fraction of users with poor experience



(a) Rebuffering Ratio



(b) Video Startup Time

Key take-away points

A "global control plane" can optimize

- Select CDN and bit rate
- Switching midstream for both of the above

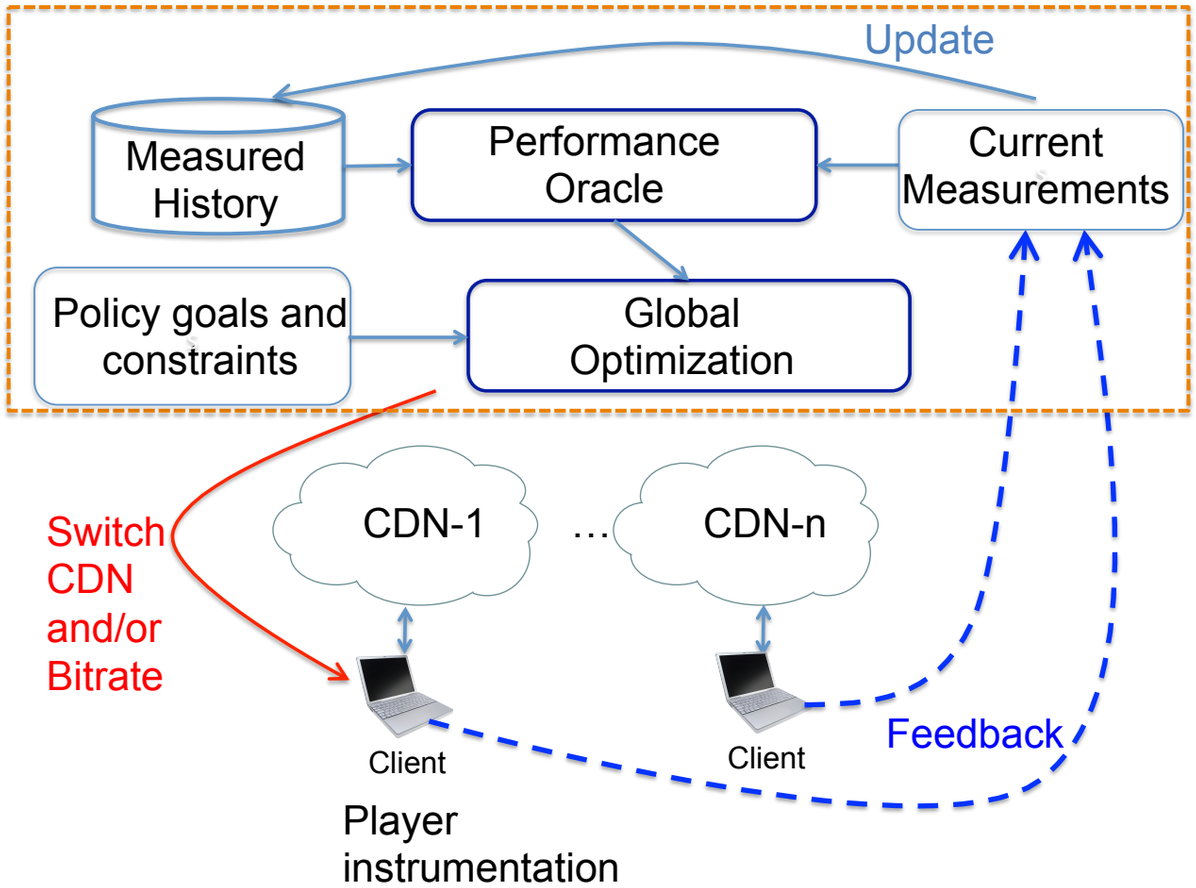


Figure 6: Overview of a video control plane

Discussion

How did they get all this data?

- 200 million viewing sessions from over 50 million viewers across 91 content providers globally

Discussion

"by 2014, video traffic will constitute more than 90% of the total traffic on the Internet"

Does this mean video is essentially the only type of traffic that's important?

Small group discussion

1. "One possible reason for such variability in the quality observed with CDNs is the load on the CDN." Other reasons?

2. Would their design work for latency-sensitive requests, like web browsing? What would be the challenges?

Announcements

Thursday: Wireless

- "Mirror Mirror on the Ceiling" [Zhou et al, SIGCOMM'12]
- Nitin Vaidya guest lecture