Respect the ORIGIN! A Best-case Evaluation of Connection Coalescing

Sudheesh Singanamalla
Muhammad Talha Paracha
Suleman Ahmad
Jonathan Hoyland
Luke Valenta
Yevgen Safronov
Peter Wu
Andrew Galloni
Vasileios Giotsas
Kurtis Heimerl
Nick Sullivan
Christopher A. Wood
Marwan Fayed

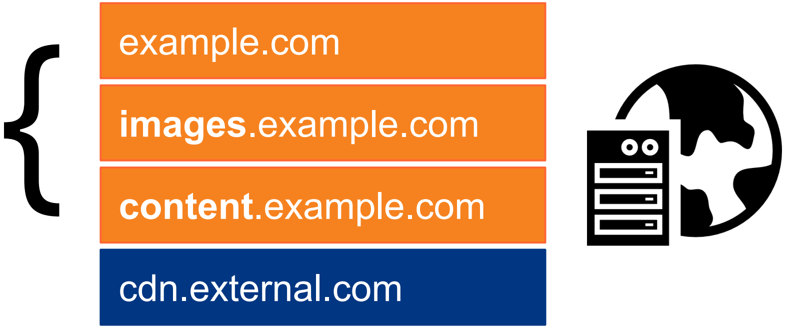
Under Submission

IMC’22

What is connection coalescing?

WANT:

Same IP addresses but results in multiple **possibly blocking** DNS queries.



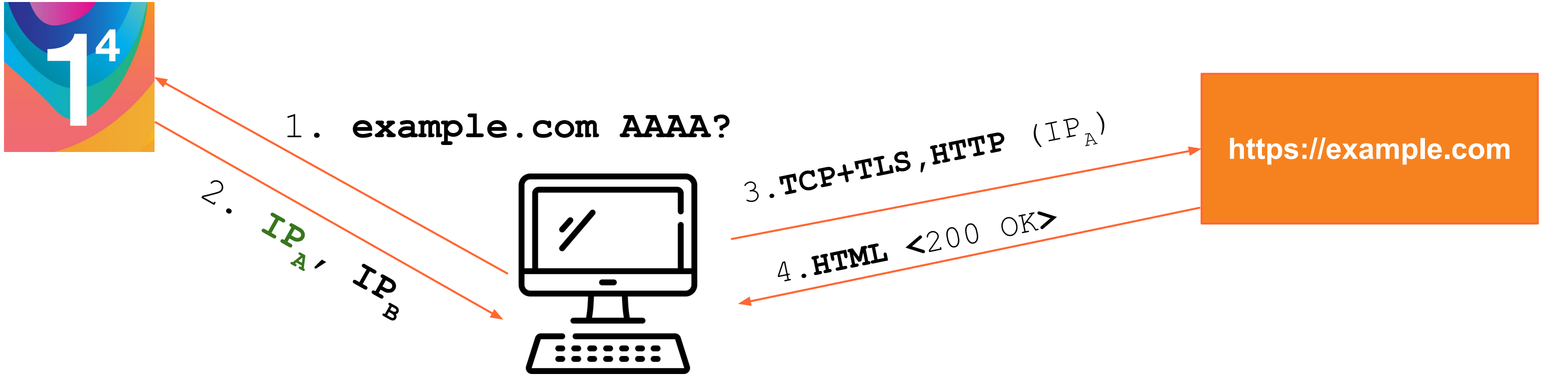
example.com

images.example.com

content.example.com

cdn.external.com

GET:



1. **example.com AAAA?**

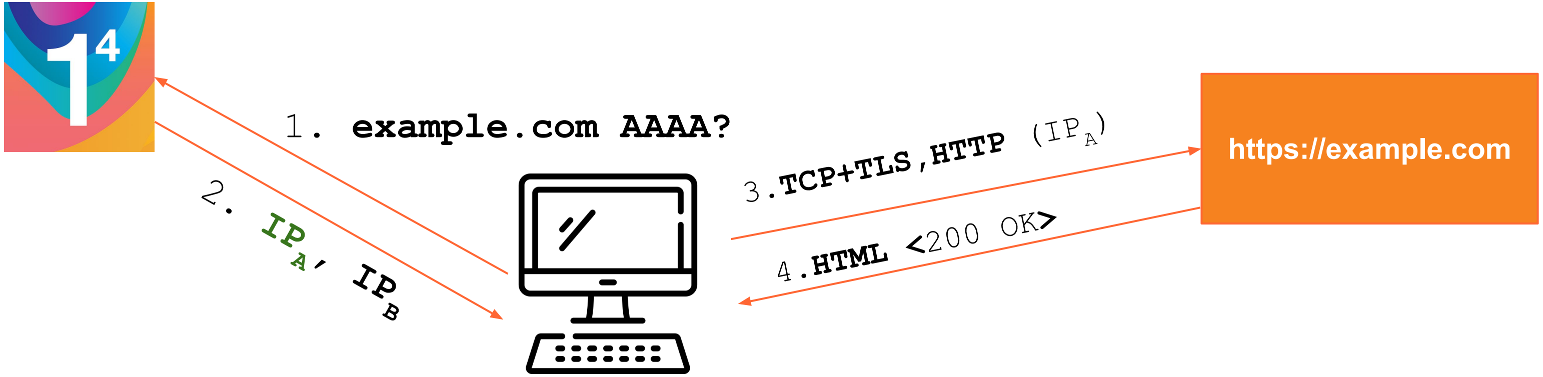
2. **IP_A, IP_B**

3. **TCP+TLS,HTTP** (IP_A)

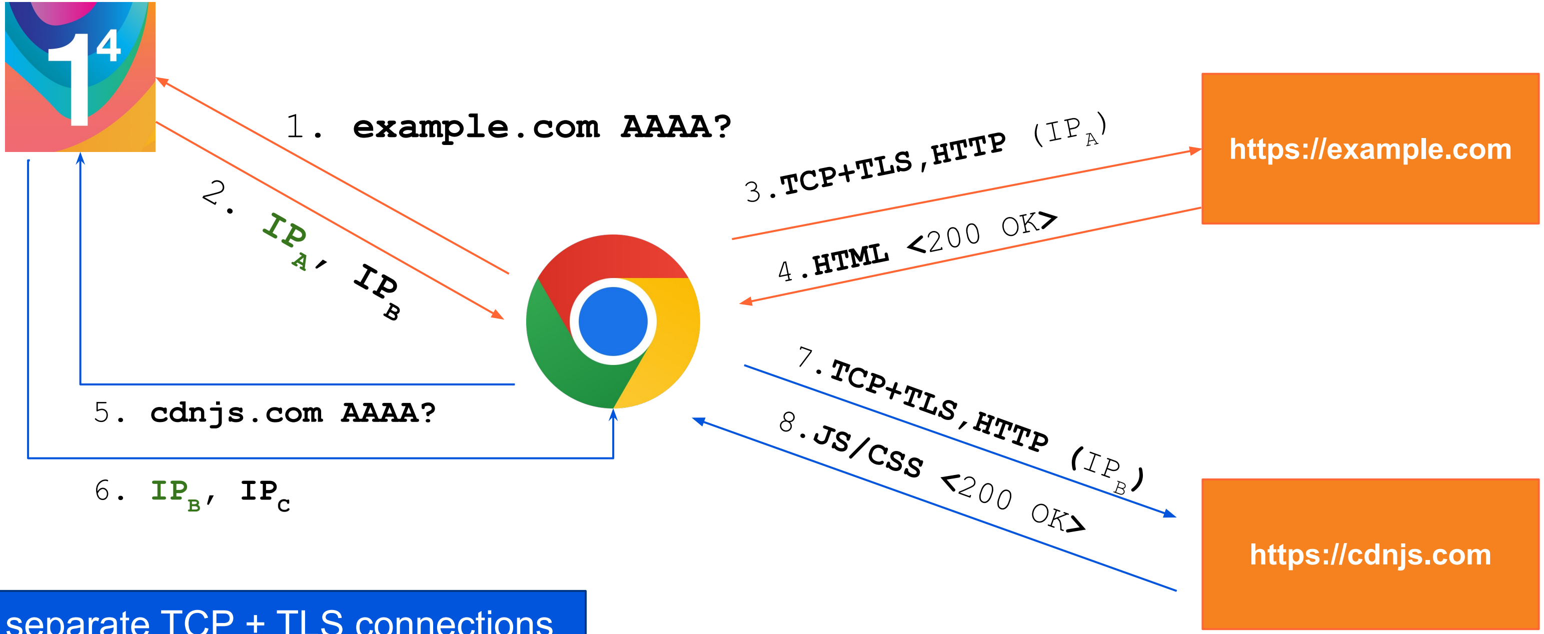
4. **HTML <**200 OK**>**

https://example.com

Next: What happens for subresources?

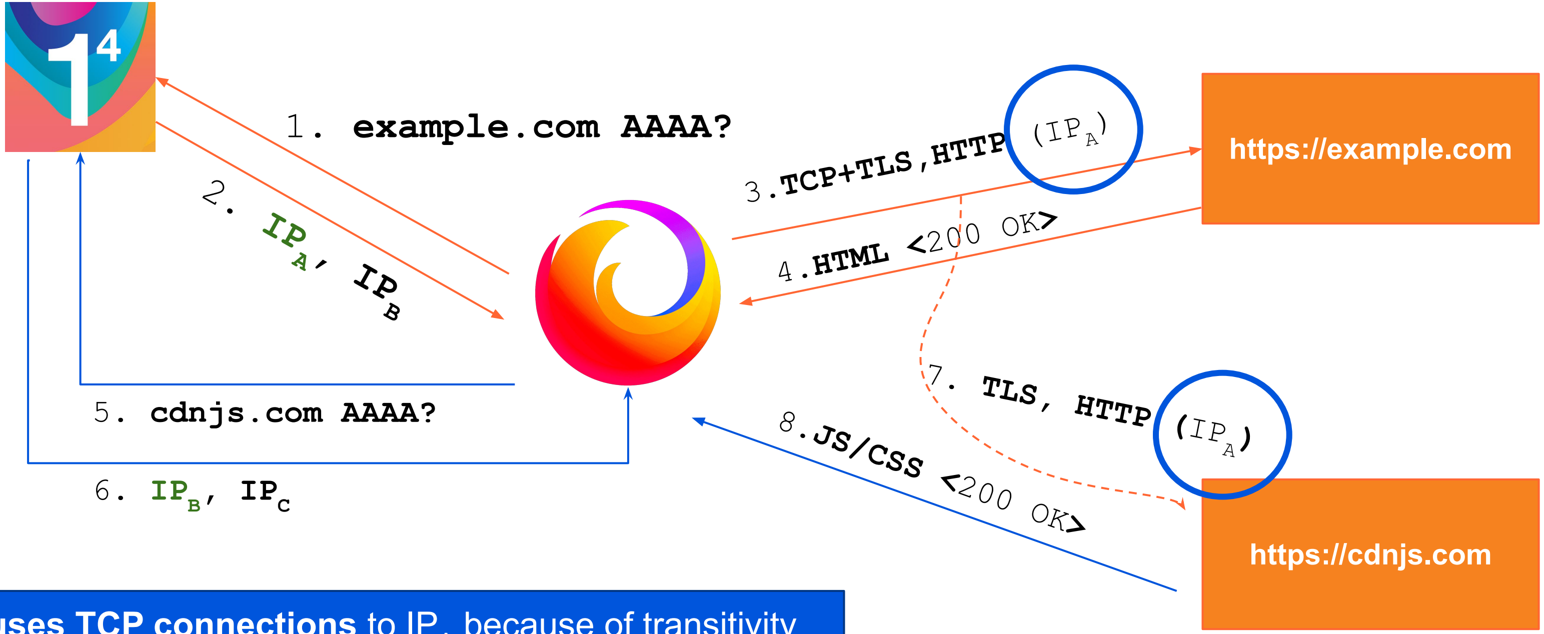


Chrome’s Approach: IP addresses for different hostnames must match



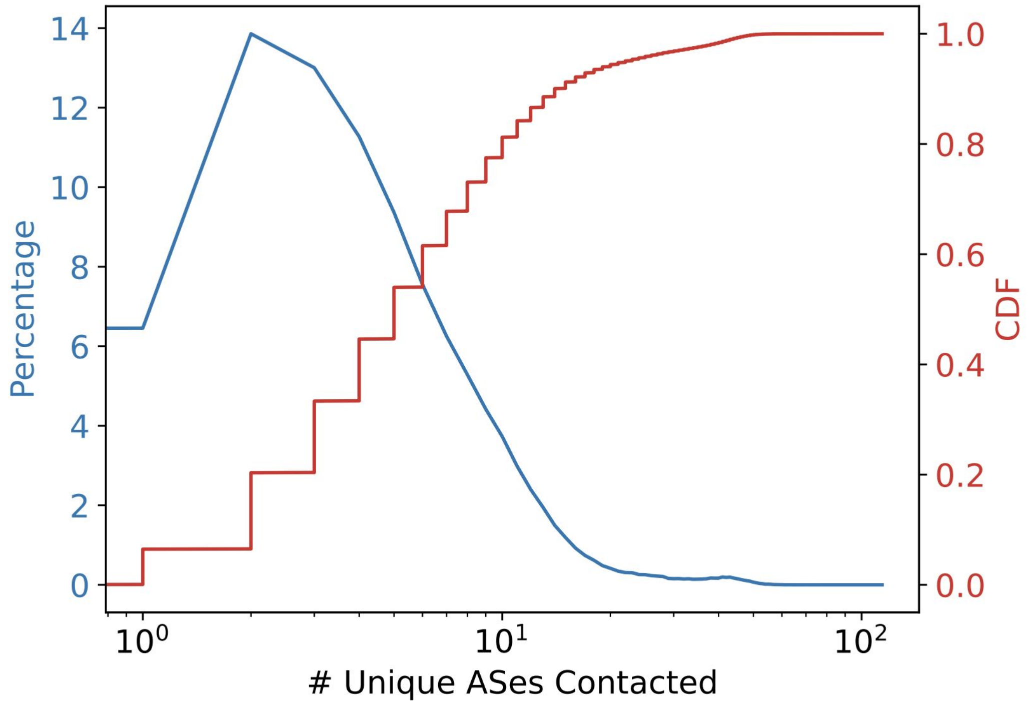
Two separate TCP + TLS connections to two different IPs (IP$_A$, IP$_B$)

Firefox's Approach: Transitivity between sets of IPs



Reuses TCP connections to IP_A because of transitivity
($IP_A \sim IP_B \sim IP_C$)

Where are the subresources located?

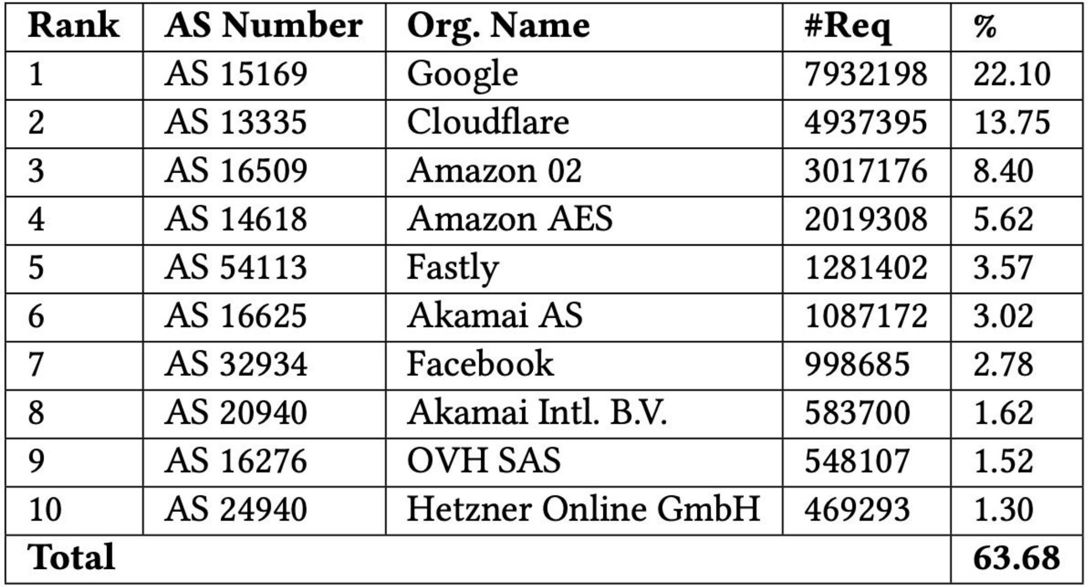


Insights:

1. 14% of web pages have a dependency on resources from one other AS.

2. More than 50% of webpages need no more than 6 ASes for all subresources.

Where are the subresources? … Coalescing favours CDNs



Rank	AS Number	Org. Name	#Req	%
1	AS 15169	Google	7932198	22.10
2	AS 13335	Cloudflare	4937395	13.75
3	AS 16509	Amazon 02	3017176	8.40
4	AS 14618	Amazon AES	2019308	5.62
5	AS 54113	Fastly	1281402	3.57
6	AS 16625	Akamai AS	1087172	3.02
7	AS 32934	Facebook	998685	2.78
8	AS 20940	Akamai Intl. B.V.	583700	1.62
9	AS 16276	OVH SAS	548107	1.52
10	AS 24940	Hetzner Online GmbH	469293	1.30
Total				**63.68**

Insights:

1. The top 10 ASes handle more than 60% of all web requests for subresources

2. Connection re-use potential (Min. number of connections) **could be approximated** to number of unique ASes contacted.

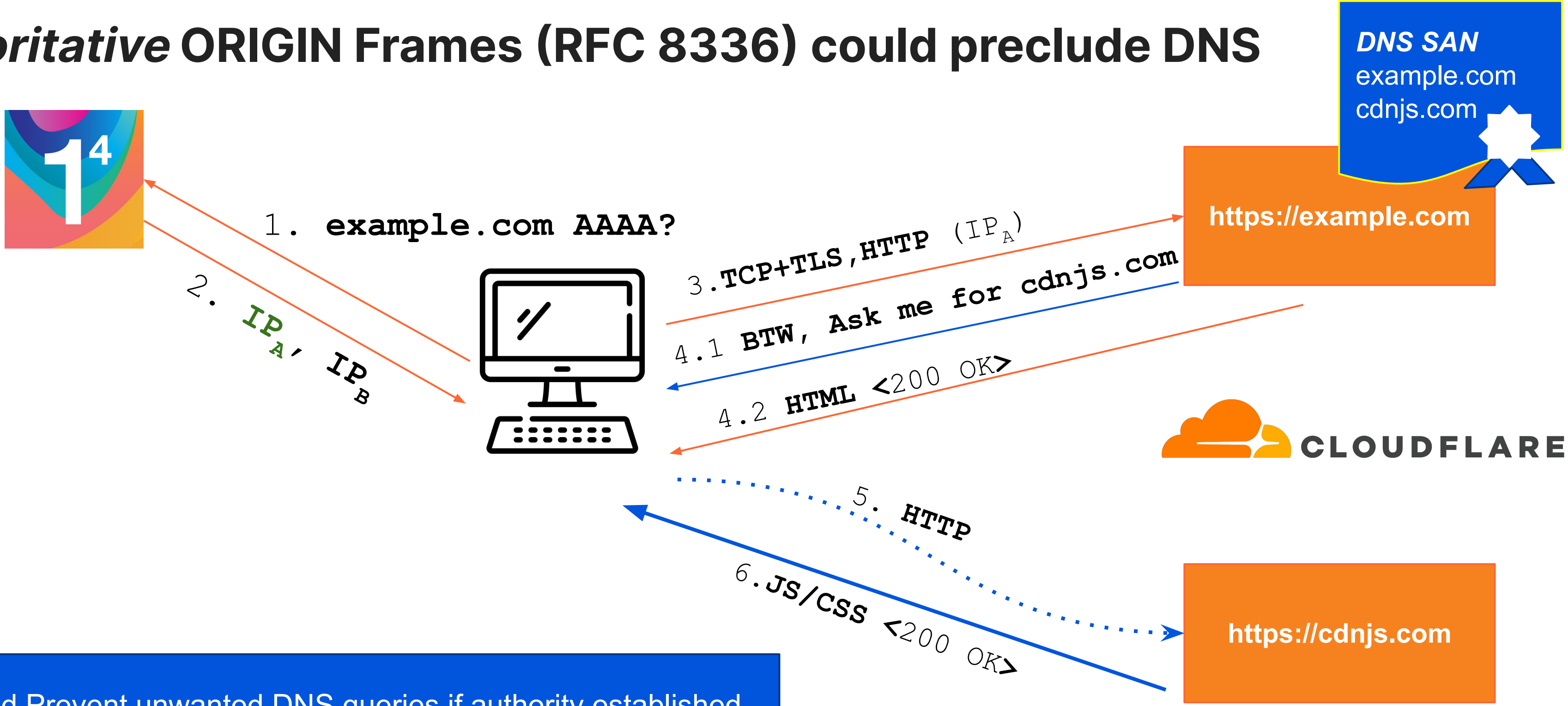
Challenges with ORIGIN Frames (RFC 8336)

1. Default ORIGIN Frame standard allows any hostname(s) to be sent by the server.

2. Clients validate the hostnames in the ORIGIN frame for authenticity
   a. Firefox is the only client which supports ORIGIN Frame
   b. Clients resolve DNS queries and retrieve TLS Certificates
      i. If the IP addresses match IP based coalescing results.
      ii. Else, new TCP+TLS connections are made.

3. Lack of server software support for ORIGIN Frames.

Authoritative ORIGIN Frames (RFC 8336) could preclude DNS



DNS SAN
example.com
cdnjs.com

1. example.com AAAA?

2. IP_A, IP_B

3. TCP+TLS, HTTP (IP_A)

4.1 BTW, Ask me for cdnjs.com

4.2 HTML <200 OK>

5. HTTP

6. JS/CSS <200 OK>

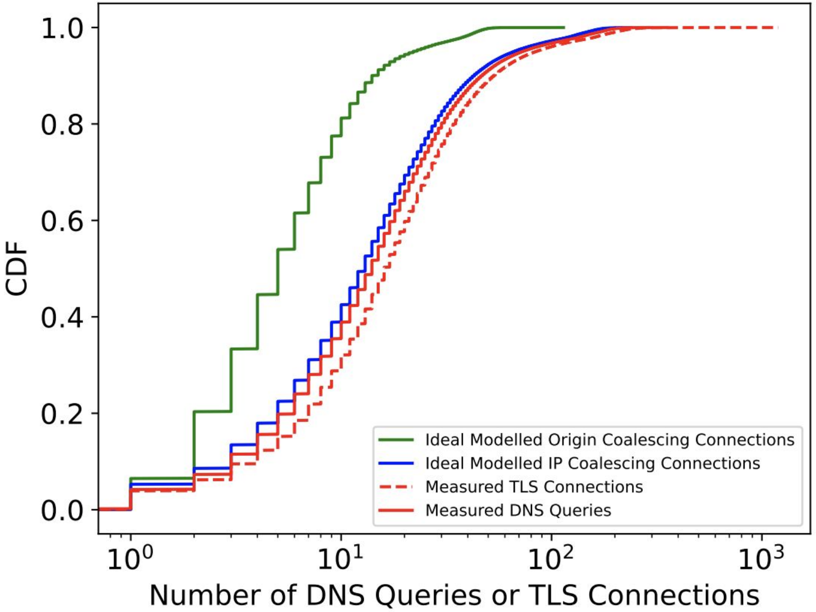
https://example.com

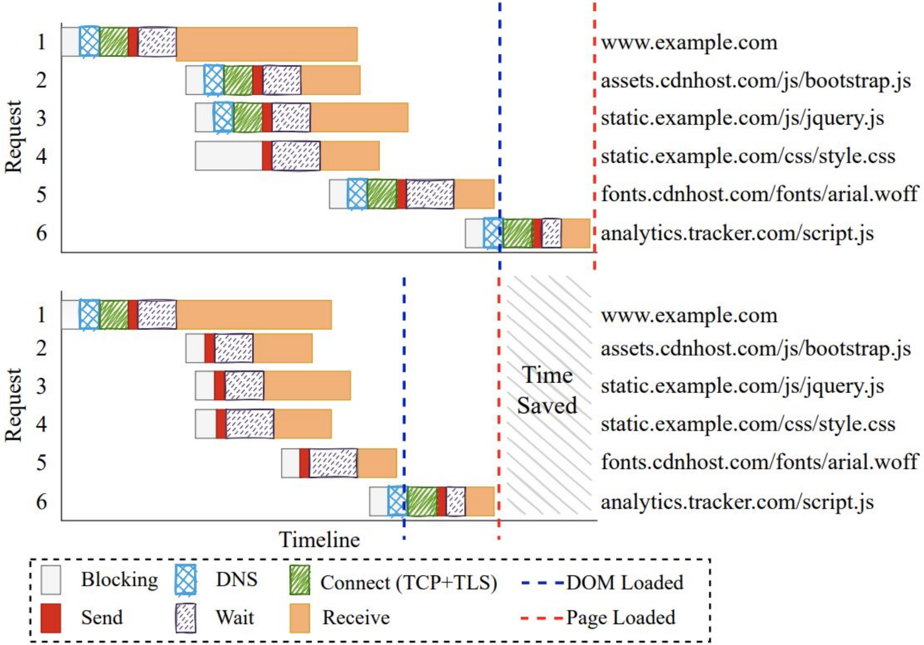
CLOUDFLARE

https://cdnjs.com

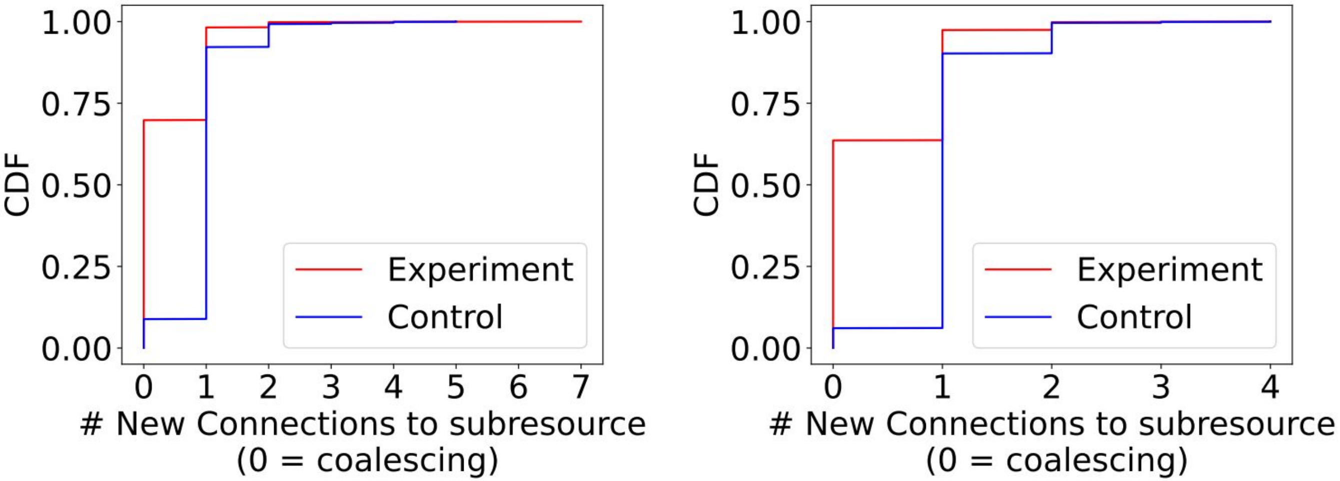
Could Prevent unwanted DNS queries if authority established

Modelling: > 60% improvement in Number of DNS and TLS connections




Request
1
2
3
4
5
6
www.example.com
assets.cdnhost.com/js/bootstrap.js
static.example.com/js/jquery.js
static.example.com/css/style.css
fonts.cdnhost.com/fonts/arial.woff
analytics.tracker.com/script.js
Time Saved
Timeline
Blocking
DNS
Connect (TCP+TLS)
DOM Loaded
Send
Wait
Receive
Page Loaded

Active Measurements: Production traffic on 5K domains

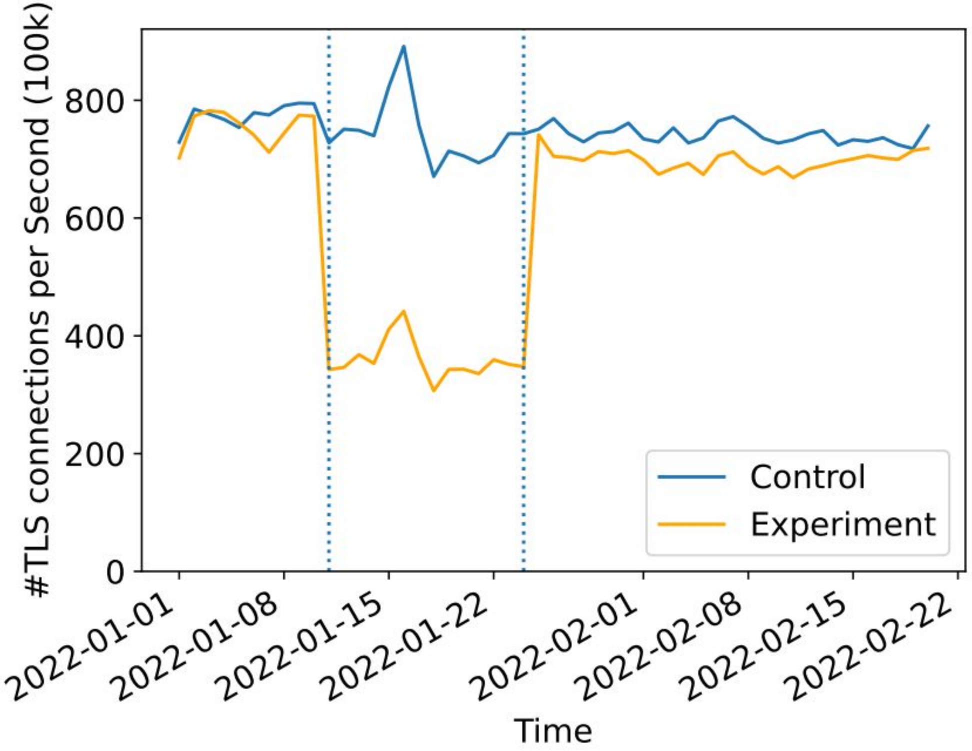


(a) IP-Based Coalescing

(b) ORIGIN Frame

1. ORIGIN Frame based coalescing approaches result in lesser overall new connections
2. Over 65% of connections can be coalesced through ORIGIN Frame (~70% IP based)

Takeaway 1: Connection Coalescing works in practice!

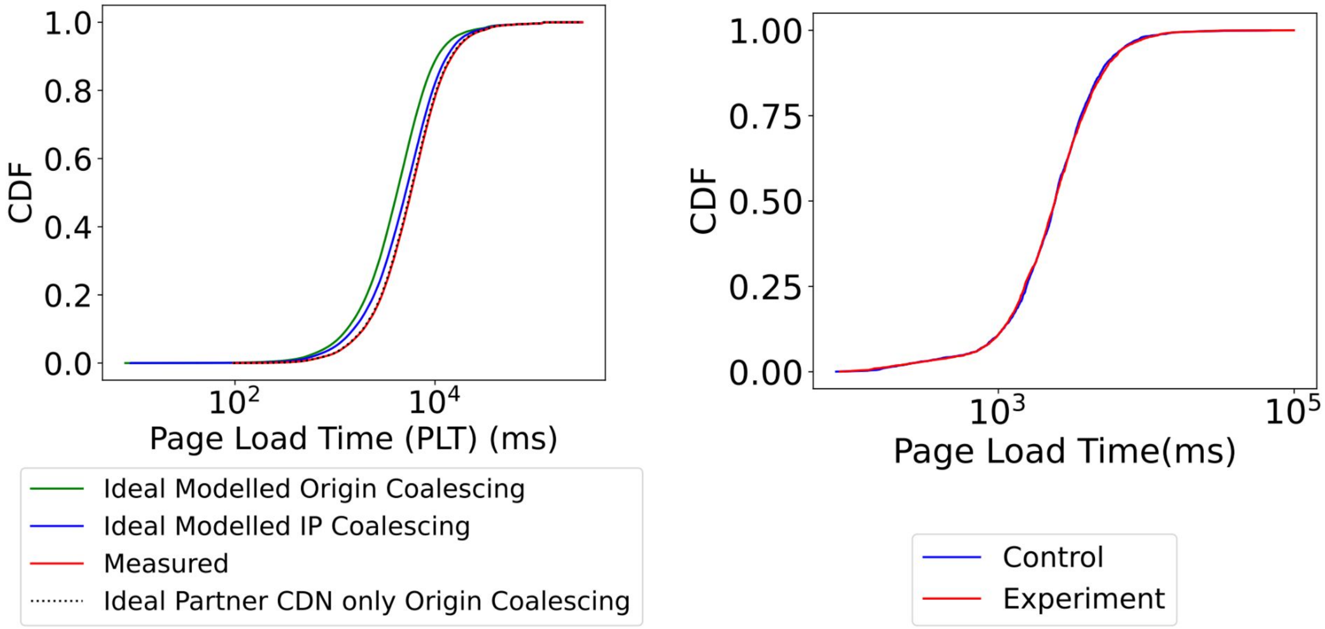


~50% reduction in number of new connections to the cdnjs hostname we attempted coalescing to.

Reduced Number of Cryptographic Certificate Validations.

Implications for reduced server compute resources.

Takeaway 2: PLT Performance is no-worse, minor improvements



(a) Measured and modelled.

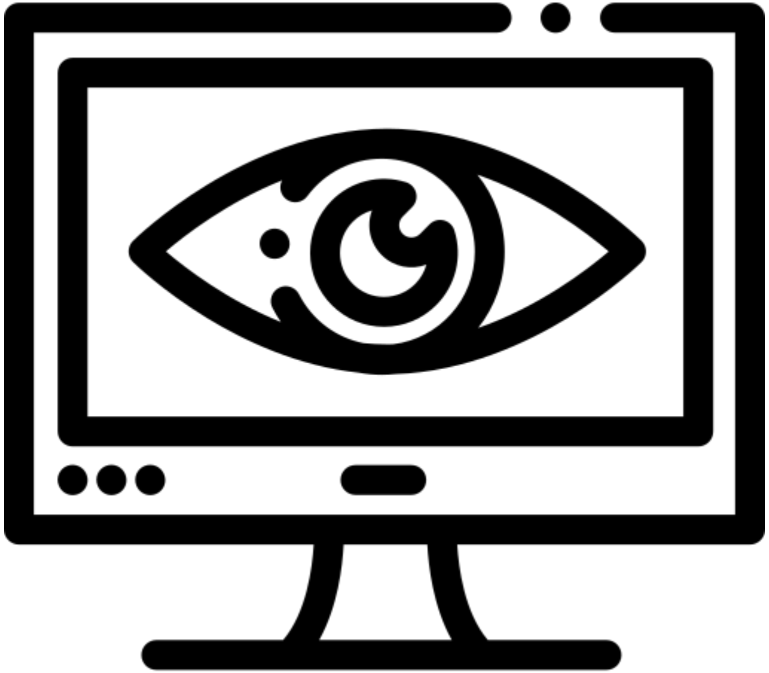
(b) IP and ORIGIN

Meaningful impact to PLT can only be seen if multiple operators enable ORIGIN frame support.

Takeaway 3: ORIGIN Frame based Coalescing improves privacy

Each coalesced connection **hides an otherwise exposed plaintext SNI** and **prevents at-least one additional plaintext DNS** query-response.

Potentially improved fingerprinting resistance but more detailed studies are needed.



Takeaway 4: Marwan wants you to know …

Connection Coalescing is ***NOT*** about performance!

Questions:

-- Unintended ripple effects in non RFC compliant HTTP/2 stacks?

-- HTTP/3? It has no ORIGIN frame equivalent!

Thank You!

sudheesh@cloudflare.com / sudheesh@cs.washington.edu
marwan@cloudflare.com