ECE458/ECE750T27: Computer Security

Web Security – Basics of Websites

Dr. Kami Vaniea,
Electrical and Computer Engineering
kami.vaniea@uwaterloo.ca

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

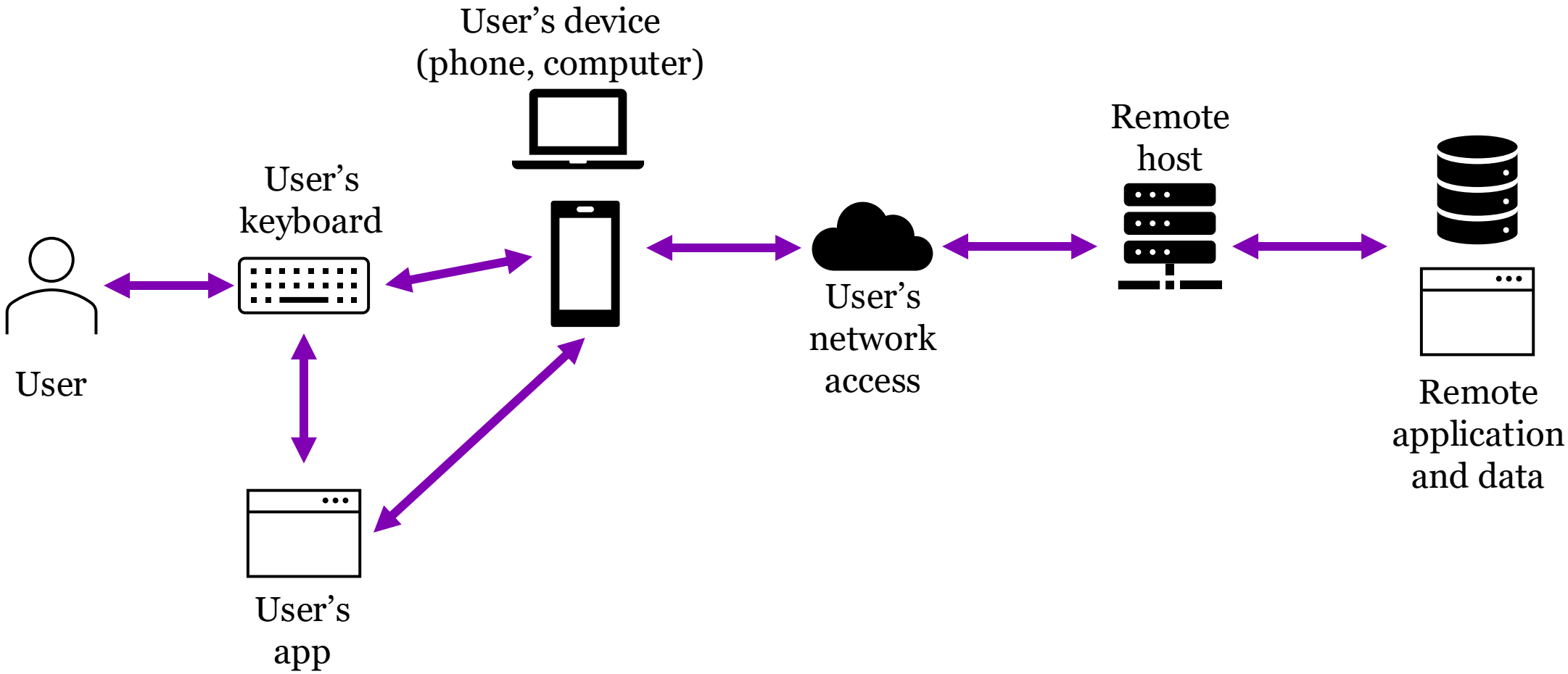
First, the news...

- First 5 minutes we talk about something interesting and recent
- You will not be tested on the news part of lecture
- You may use news as an example on tests
- Why do this?
    1. Some students show up late for various good reasons
    2. Reward students who show up on time
    3. Important to see real world examples

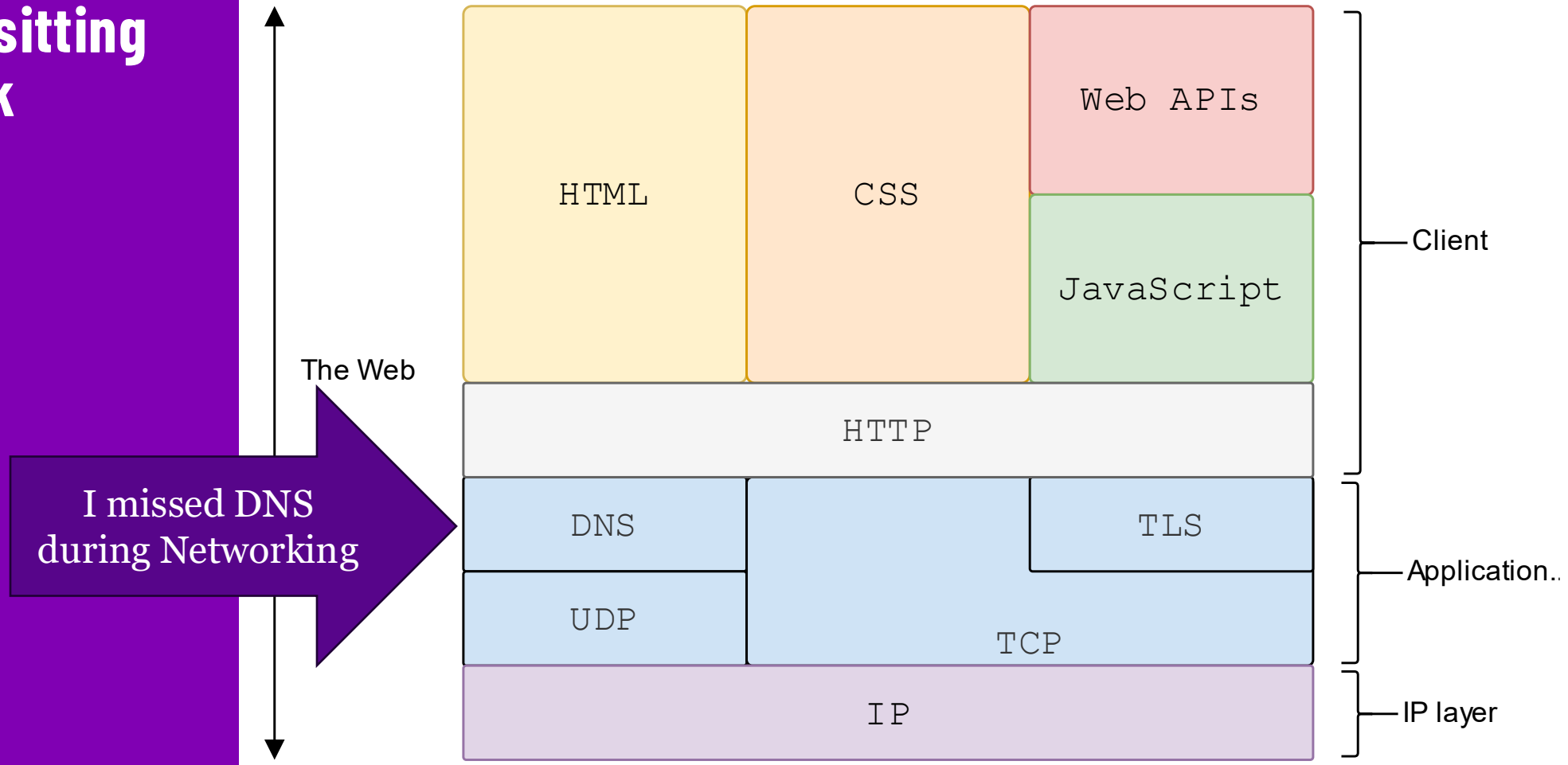
Web Security

- Today: How the internet works
  - DNS
  - Website dynamic construction
  - Data typically sent and logged
  - Cookies
  - Attacks on privacy - trackers
- Next lecture: Getting the user or server to execute attacker code
  - Cross Site Scripting (XSS)
  - Cross-site Request Forgery (CSRF)
  - Server-side Scripting
  - Session Hijacking

User-to-remote application vulnerability



The web is basically another layer sitting on the network model.



https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview

DOMAIN NAME SYSTEM (DNS)

Domain Name System (DNS)

- The DNS service translates human friendly URLs such as http://vaniea.com to their IP address such as 69.163.145.230.
- Mappings between URLs and IPs are not static.
- DNS servers keep track of current mappings.
- One domain, such as google.com, may have many IP addresses associated with it since several computers are used to host the domain

```
1003 > dig google.ca

; <<>> DiG 9.18.30-0ubuntu0.24.04.2-Ubuntu <<>> google.ca
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 42536
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 4, ADDITIONAL: 9

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags:; udp: 1220
; COOKIE: 472addeb1974fd3d0100000068703137bd7eb011b4322f8d (good)
;; QUESTION SECTION:
;google.ca.                     IN      A

;; ANSWER SECTION:
google.ca.              292     IN      A       142.251.32.67

;; AUTHORITY SECTION:
google.ca.              60527   IN      NS      ns3.google.com.
google.ca.              60527   IN      NS      ns2.google.com.
google.ca.              60527   IN      NS      ns1.google.com.
google.ca.              60527   IN      NS      ns4.google.com.

;; ADDITIONAL SECTION:
ns1.google.com.         15998   IN      A       216.239.32.10
ns2.google.com.         19125   IN      A       216.239.34.10
ns3.google.com.         74037   IN      A       216.239.36.10
ns4.google.com.         70510   IN      A       216.239.38.10
ns1.google.com.         37624   IN      AAAA    2001:4860:4802:32::a
ns2.google.com.         74502   IN      AAAA    2001:4860:4802:34::a
ns3.google.com.         79559   IN      AAAA    2001:4860:4802:36::a
ns4.google.com.         60465   IN      AAAA    2001:4860:4802:38::a

;; Query time: 18 msec
;; SERVER: 129.97.2.1#53(129.97.2.1) (UDP)
;; WHEN: Thu Jul 10 17:31:35 EDT 2025
;; MSG SIZE  rcvd: 340
```

```
kvaniea@brendel:~$
1247 > host facebook.com
facebook.com has address 31.13.77.36
facebook.com has IPv6 address 2a03:2880:f101:83:face:b00c:0:25de
facebook.com mail is handled by 10 msgin.vvv.facebook.com.
kvaniea@brendel:~$
1248 > host microsoft.com
microsoft.com has address 23.96.52.53
microsoft.com has address 104.40.211.35
microsoft.com has address 104.43.195.251
microsoft.com has address 191.239.213.197
microsoft.com has address 23.100.122.175
microsoft.com mail is handled by 10 microsoft-com.mail.protection.outlook.com.
kvaniea@brendel:~$
1249 > host google.com
google.com has address 172.217.23.14
google.com has IPv6 address 2a00:1450:4009:801::200e
google.com mail is handled by 50 alt4.aspmx.l.google.com.
google.com mail is handled by 20 alt1.aspmx.l.google.com.
google.com mail is handled by 10 aspmx.l.google.com.
google.com mail is handled by 30 alt2.aspmx.l.google.com.
google.com mail is handled by 40 alt3.aspmx.l.google.com.
kvaniea@brendel:~$
1250 > host vaniea.com
vaniea.com has address 64.90.44.164
vaniea.com mail is handled by 0 vade-in2.mail.dreamhost.com.
vaniea.com mail is handled by 0 vade-in1.mail.dreamhost.com.
```

```
kvaniea@brendel:~$
1253 > host -v facebook.com
Trying "facebook.com"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 29731
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 2, ADDITIONAL: 4

;; QUESTION SECTION:
;facebook.com.                  IN      A

;; ANSWER SECTION:
facebook.com.           300     IN      A       157.240.2.35

;; AUTHORITY SECTION:
facebook.com.           172261  IN      NS      b.ns.facebook.com.
facebook.com.           172261  IN      NS      a.ns.facebook.com.

;; ADDITIONAL SECTION:
a.ns.facebook.com.      172261  IN      A       69.171.239.12
a.ns.facebook.com.      172261  IN      AAAA    2a03:2880:fffe:c:face:b00c:0:35
b.ns.facebook.com.      172261  IN      A       69.171.255.12
b.ns.facebook.com.      172261  IN      AAAA    2a03:2880:ffff:c:face:b00c:0:35
```

Name Servers (NS) for Facebook.com
These servers notify other computers what Facebook's current IP is

IANA specifies the root name servers for DNS.

Domain Names

Overview
Root Zone Management
- Overview
- Root Database
- Hint and Zone Files
- Change Requests
- Instructions & Guides
- **Root Servers**

.INT Registry
.ARPA Registry
IDN Practices Repository
Root Key Signing Key (DNSSEC)
Reserved Domains

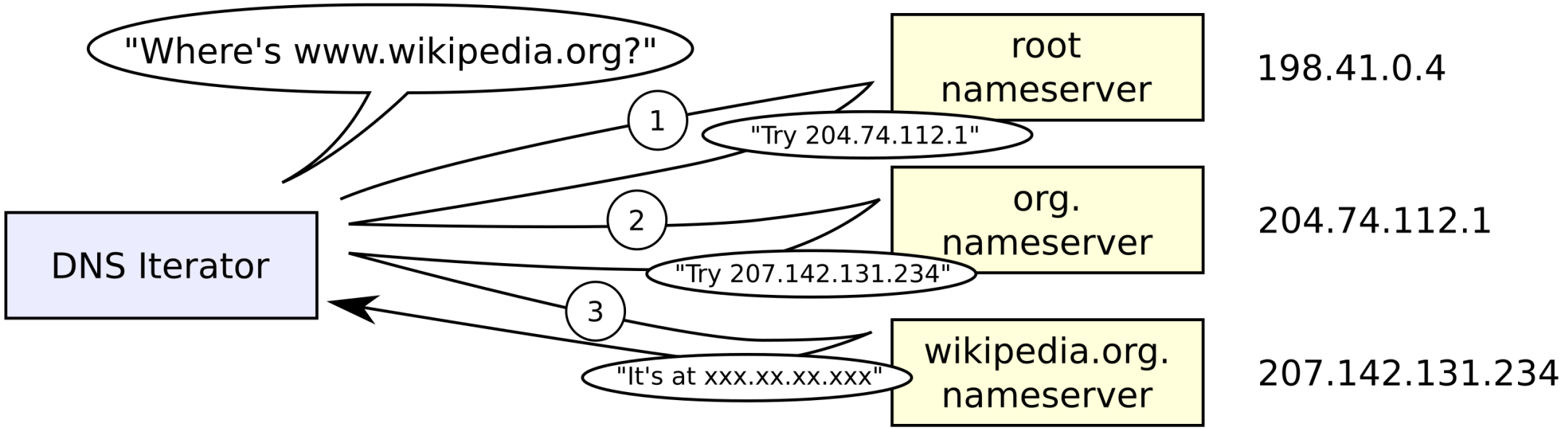
Root Servers

The authoritative name servers that serve the DNS root zone, commonly known as the "root servers", are a network of hundreds of servers in many countries around the world. They are configured in the DNS root zone as 13 named authorities, as follows.

List of Root Servers

HOSTNAME	IP ADDRESSES	OPERATOR
a.root-servers.net	198.41.0.4, 2001:503:ba3e::2:30	Verisign, Inc.
b.root-servers.net	170.247.170.2, 2801:1b8:10::b	University of Southern California, Information Sciences Institute
c.root-servers.net	192.33.4.12, 2001:500:2::c	Cogent Communications
d.root-servers.net	199.7.91.13, 2001:500:2d::d	University of Maryland
e.root-servers.net	192.203.230.10, 2001:500:a8::e	NASA (Ames Research Center)
f.root-servers.net	192.5.5.241, 2001:500:2f::f	Internet Systems Consortium, Inc.
g.root-servers.net	192.112.36.4, 2001:500:12::d0d	US Department of Defense (NIC)
h.root-servers.net	198.97.190.53, 2001:500:1::53	US Army (Research Lab)
i.root-servers.net	192.36.148.17, 2001:7fe::53	Netnod
j.root-servers.net	192.58.128.30, 2001:503:c27::2:30	Verisign, Inc.
k.root-servers.net	193.0.14.129, 2001:7fd::1	RIPE NCC
l.root-servers.net	199.7.83.42, 2001:500:9f::42	ICANN

DNS request happens before a webserver can be contacted Because the OS/Browser does not know where the server is



By Lion Kimbro - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=386517

Local DNS cache

- DNS is cached at several levels
- When the Operating System resolver does a lookup, it keeps the record it found cached locally
- Browsers can also do the resolution on their own and they also cache

```
google.com
----------------------------------------
Record Name . . . . . : google.com
Record Type . . . . . : 1
Time To Live  . . . . : 137399
Data Length . . . . . : 4
Section . . . . . . . : Answer
A (Host) Record . . . : 142.251.41.46
```

How to contact the name server?

- Security was not initially a concern
  - DNS sent over UDP – no encryption at all
  - The initial solutions only addressed Integrity, not Confidentiality
- Encrypting DNS connections is just now becoming common

- 1989 – DNS Introduced
- 2004 – DNS Security Extensions (DNSSEC) - added signatures
- 2011 - DNSCrypt protocol
- 2016 – DNS over TLS
- 2018 – DNS over HTTPS
- 2019 – Chrome uses DNS over HTTPS by default (US only)
- 2020 – Firefox uses DNS over HTTPS by default

Think-pair-share

- How could unencrypted DNS be used to cause a user to visit the wrong web server?
- What technology have we already learned about that can be used to detect the above attack?
- What attack on privacy becomes possible with unencrypted DNS?

DNS Servers are soft targets for attackers, take out the mapping and the website goes "offline"

DoS attack on major DNS provider brings Internet to morning crawl [Updated]

Dyn's US East region hit hardest in attack that affected Twitter, Reddit.

SEAN GALLAGHER - OCT 21, 2016 1:59 PM UTC

118

Update (12:04p ET): A second wave of DDoS attacks against Dyn is underway, as of noon Eastern Time today. Dyn is continuing to work on the issue. Our original story follows below; further updates will be added as information becomes available.

A distributed denial of service attack against Dyn, the dynamic DNS service, affected the availability of dozens of major websites and Internet services this morning, including Twitter and Reddit. The attack, which began this morning at 7:10am Eastern Time (12:10pm UK), is apparently focused on Dyn's US East Coast name servers.

"This morning, Dyn received a global DDoS attack on our Managed DNS infrastructure in the east coast of the United States," Doug Madory, Director of Internet Analysis at Dyn, said in an e-mail sent to Ars this morning. "DNS traffic resolved from east coast name server locations are experiencing a service interruption during this time." By 9:20am ET this morning, Dyn had mitigated the attack and services returned to normal.

[Update, 1:20 PM ET] Less than three hours later, the attack began again, and is still in progress.

Syrian group cited as New York Times outage continues

By Heather Kelly, CNN

Updated 1330 GMT (2130 HKT) August 29, 2013

The hackers gained access to a Melbourne IT reseller account using a phishing email and proceeded to change the DNS records of multiple domains, including NYTimes.com, according to the company.

The N

Stor

New
probl

Syria
York

The group is loyal to Syrian President Bashar Al-Assad

Twitter also experienced problems on Tuesday due to a similar attack

multiple attacks on media websites in recent months and, on Twitter, took credit for a sophisticated hack that had hobbled the Times' news site for roughly 20 hours.

"The @nytimes attack was going to deliver an anti-war message but our server couldn't last for 3 minutes," the group posted on its Twitter feed at about 9:40 Wednesday morning.

Facebook goes offline

- When Facebook stopped advertising their DNS Nameservers on BGP, then other websites stopped being able to lookup the IP address of facebook.com
- That error then propagated across the internet taking Facebook offline worldwide

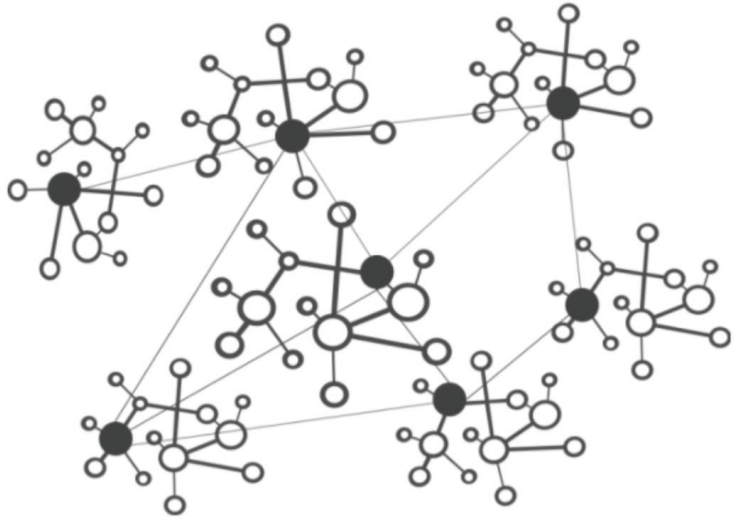
Understanding how Facebook disappeared from the Internet

10/04/2021

Celso Martinho Tom Strickx

7 min read

This post is also available in 简体中文, 繁體中文, 日本語, 한국어, Deutsch, Français, Español, Português, Русский, and Italiano.



The Internet - A Network of Networks

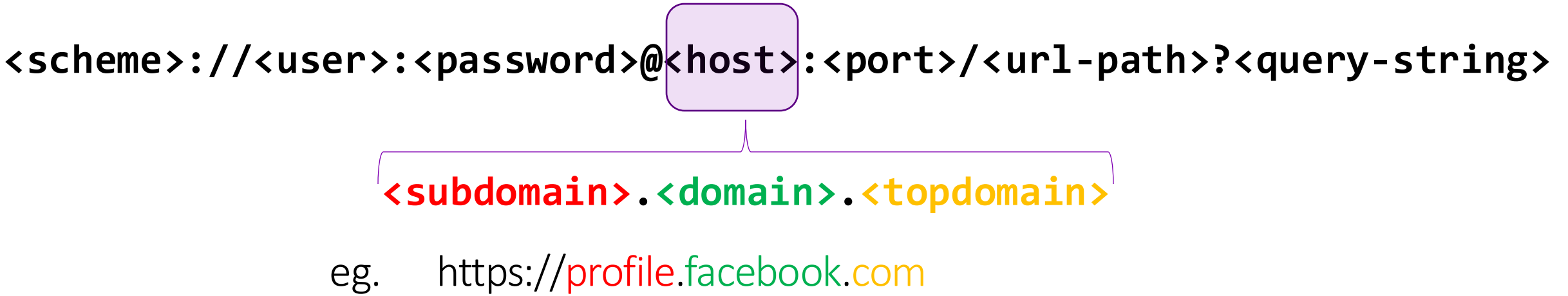
"*Facebook can't be down, can it?*", we thought, for a second.

Today at 15:51 UTC, we opened an internal incident entitled "Facebook DNS lookup returning SERVFAIL" because we were worried that something was wrong with our

URL READING

What is a URL?

- Uniform Resource Locators (URLs) are a standardized format for describing the location and access method of resources via the internet.



What is a URL?

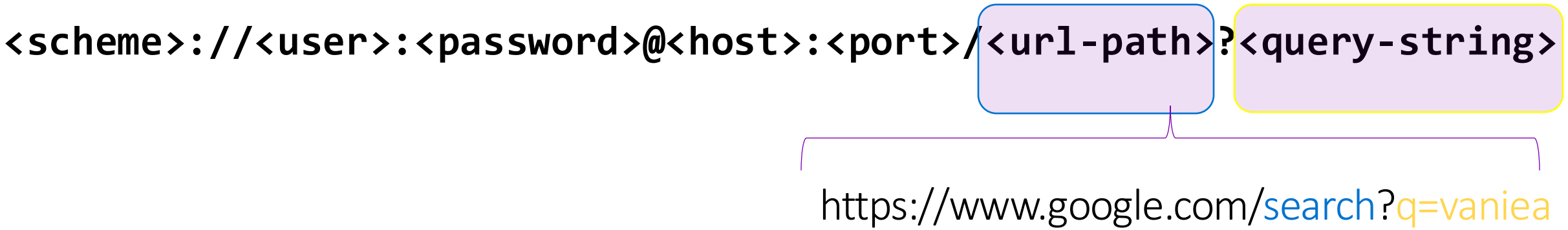
- Uniform Resource Locators (URLs) are a standardized format for describing the location and access method of resources via the internet.

What is a URL?

- Uniform Resource Locators (URLs) are a standardized format for describing the location and access method of resources via the internet.

What is a URL?

- Uniform Resource Locators (URLs) are a standardized format for describing the location and access method of resources via the internet.



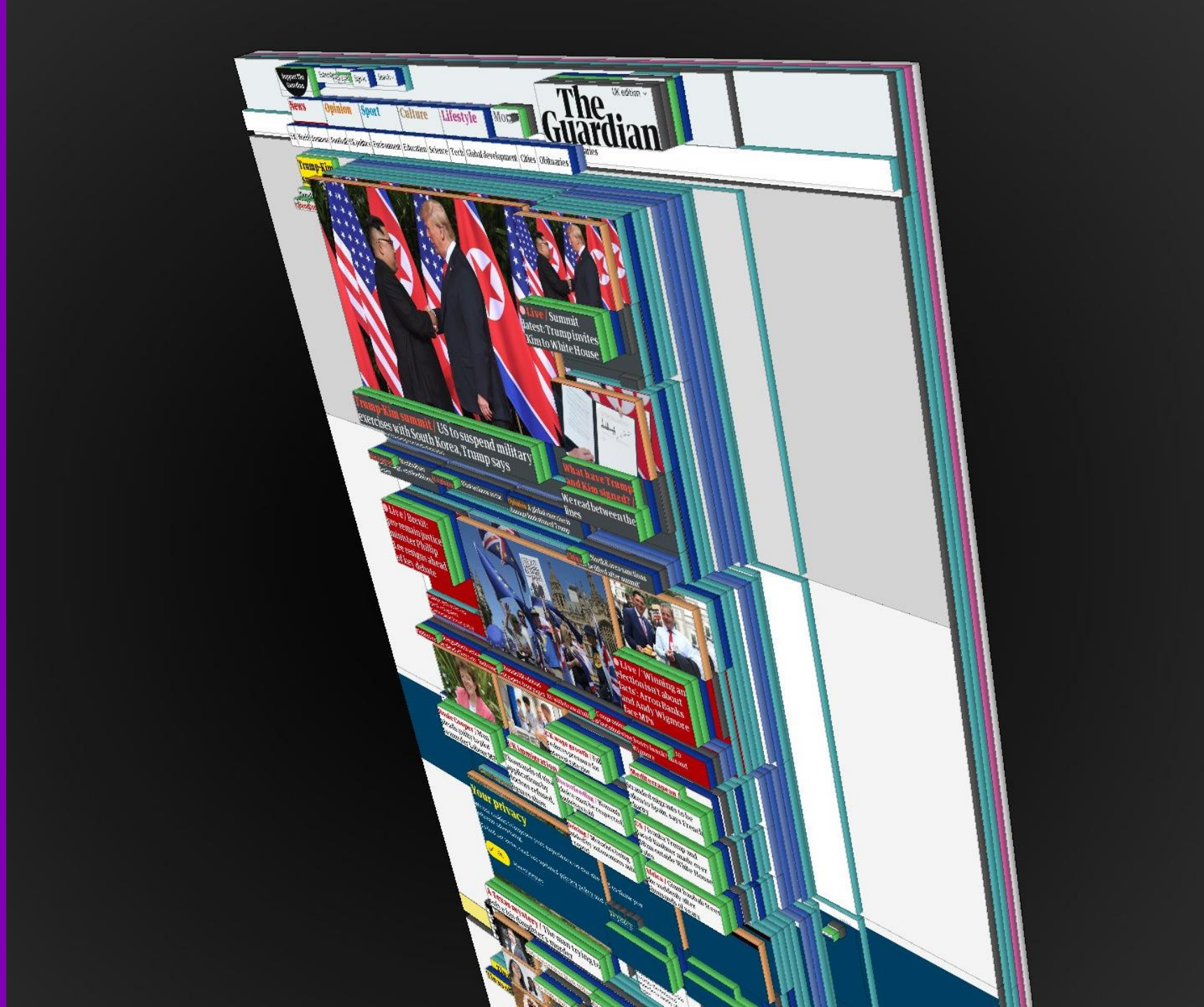
HOW WEB PAGES ARE BUILT

Once you click a link, your web browser will attempt to load the page.

Lets look at what the browser is doing.

Web pages are not flat. Nor are they read-only.

They are dynamically constructed out of a combination of media provided by the main website and by 3rd party websites.



Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

Think-pair-share

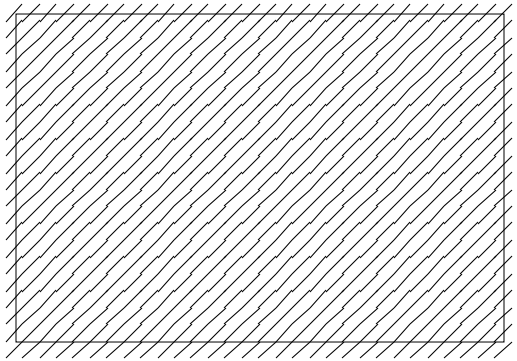
- Where are the trust boundaries in this website loading?

Imagine an overly simple (fake) website: cutedogs.com

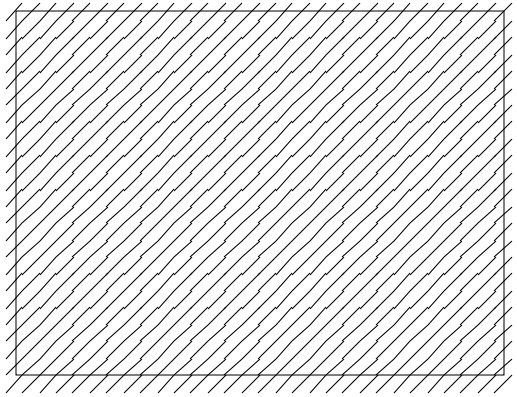
```
<html>
        <head>
                <title>Cute Dogs!</title>
                <link href="./style.css" rel="stylesheet">
        </head>
        <body>
                <img href="instagram.com/dog.jpg"> Cute dog …
                <img href="instagram.com/sad_dog.jpg"> Play …
                <img href="doubleclick.com/invisible.jpg"
                        width="0" height="0">
                <script>
                <!- Code here that loads
                    https://beacon.krxd.net/tracker.js
                   https://connect.facebook.net/connect.js
                -->
                </script>

        </body>
</html>
```

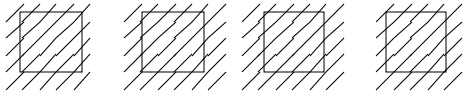
Cute Dogs!



Cute dog sleeping on a park bench.



Playing with a ball all day is hard work. But that is no reason to release the ball.

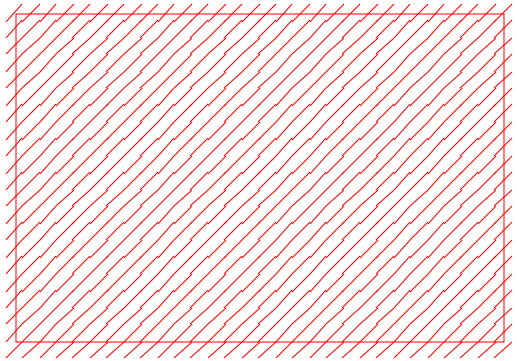


First, the computer fetches the main page. This page is similar to a template, or a recipe in that it contains instructions on how to build the rest of the page.

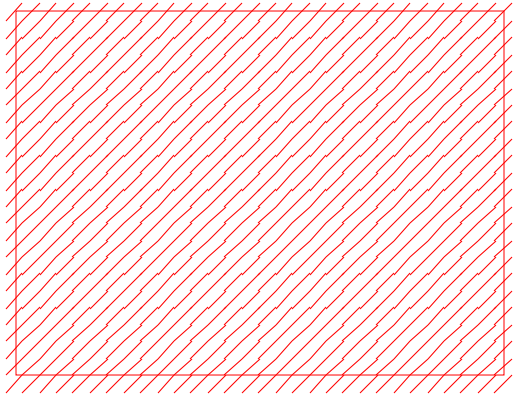
Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!



Cute dog sleeping on a park bench.



Playing with a ball all day is hard work. But that is no reason to release the ball.

Red shaded boxes are parts of the template that are waiting to be

Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

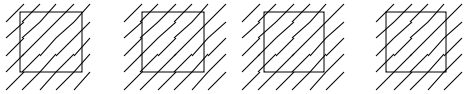
Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.



Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net

Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net
- logo.jpg from connect.facebook.net

Cute Dogs!

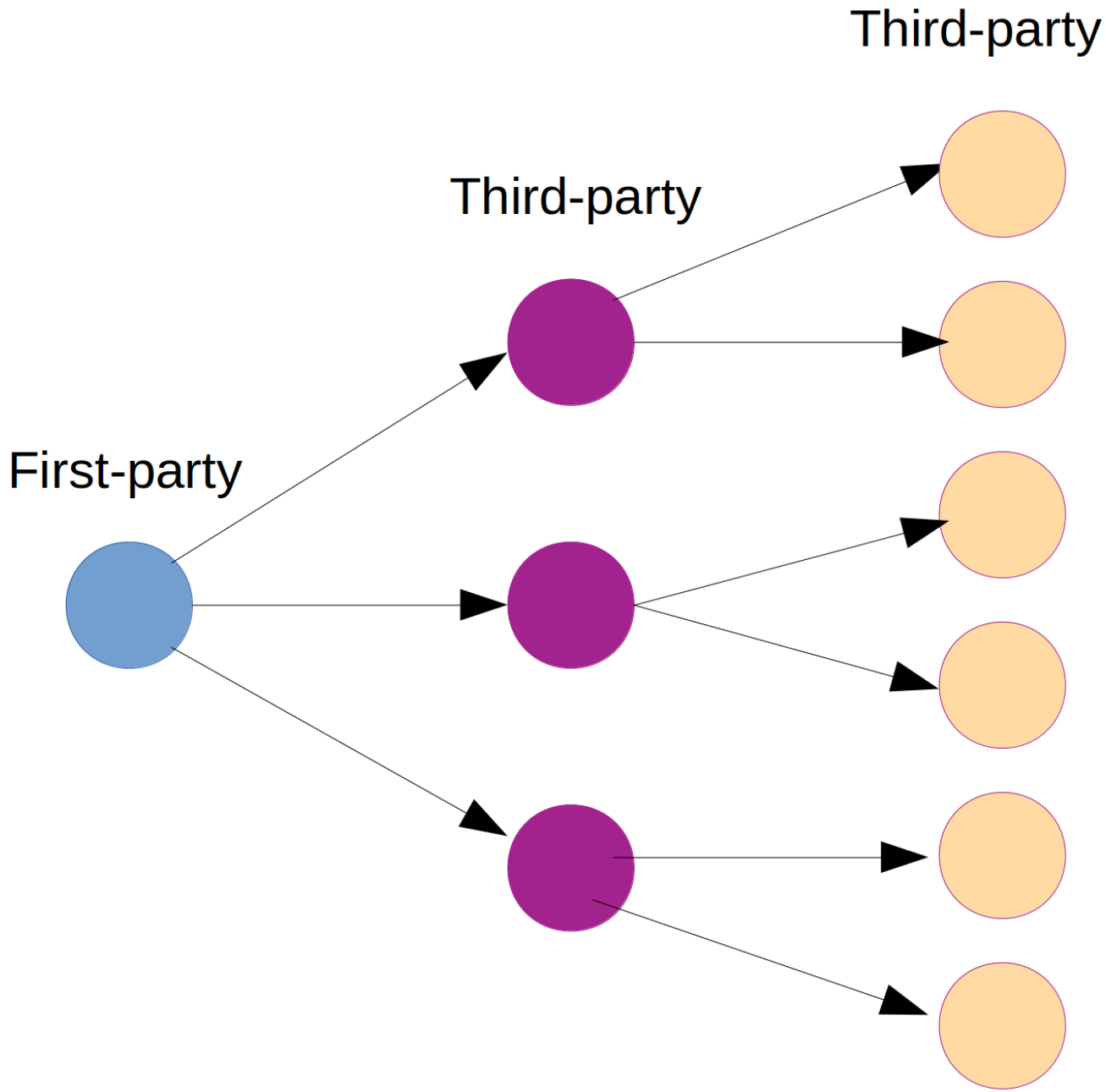
Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

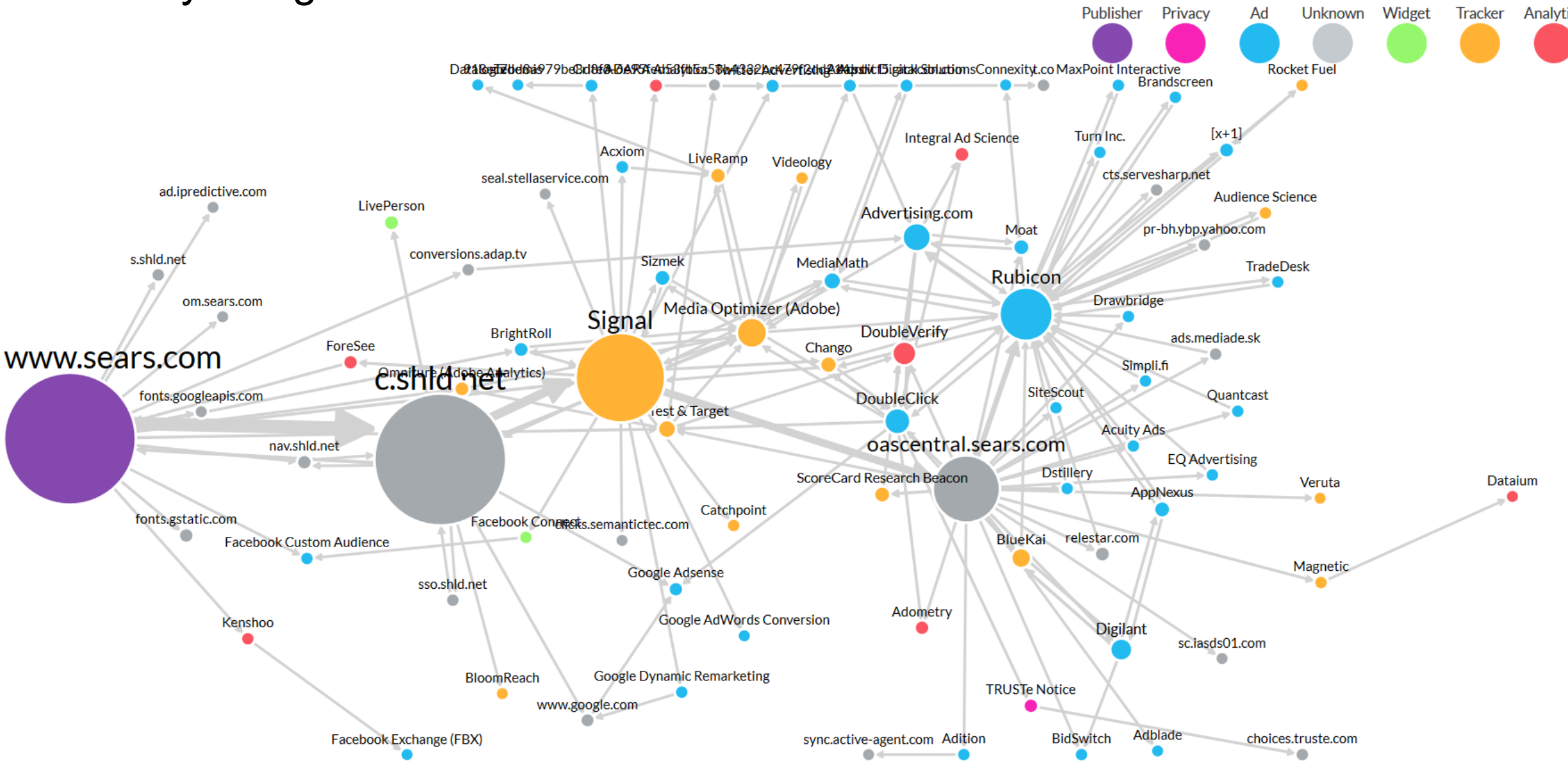
Loaded:

- index.html from cutedog.com
- dog.jpg from instagram.com
- sad_dog.jpg from instagram.com
- style.css from cutedog.com
- tracker.js from beacon.krxd.net
- invisible.jpg from doubleclick.com
- connect.js from connect.facebook.net
- logo.jpg from connect.facebook.net

Each piece of content can add new content to the page.



Ghostery: Page content sources



Headlines

Monday
11 June 2018

Edinburgh

Now 13°C

17:00	20:00	23:00
16°C	14°C	12°C

Brexit / 'Catastrophic' no deal a real possibility, says Vince Cable

Lib Dem leader urges Theresa May to stand up to hardline Brexiters in her own party

Jacob Rees-Mogg 'No need for Dover customs checks under no deal'

Theresa May PM calls for party unity over key Brexit votes

EU withdrawal bill Full list of proposed amendments

Theresa May / PM calls her response to Grenfell fire 'not good enough'

Live May admits rehousing survivors has taken 'too long'

Grenfell survivors Council added to suffering, report says

UK manufacturing / Largest monthly fall in output for five years

Arron Banks / UK

The Guardian Content Sources

- assets.guim.co.uk
- i.guim.co.uk
- interactive.guim.co.uk
- ipv4.guim.co.uk
- ipv6.guim.co.uk
- poly_x001c_ll.g
- uim.co.uk
  uploads.guim.co.uk
- api.nextgen.guardianapps.co.uk
- ophan.theguardian.com
- static.theguardian.com
- ad.360yield.com
- adservice.google.com
- adservice.google.co.uk
- apex.go.sonobi.com
- as-sec.casalemedia.com
- beacon.gu-web.net
- beacon.krxd.net
- cdn.adsafeprotected.com
- cdn-gl.imrworldwide.com
- cdn.krxd.net
- connect.facebook.net
- consumer.krxd.net
- dt.adsafeprotected.com
- dual.guim.co.uk
- elb.the-ozone-project.com
- googleads.g.doubleclick.net
- i.ytimg.com
- pagead2.googlesyndication.com
- partner.mediawallahscript.com
- pixel.adsafeprotected.com
- prebid.adnxs.com
- sb.scorecardresearch.com
- secure-au.imrworldwide.com
- secure-dcr.imrworldwide.com
- secure-gl.imrworldwide.com
- securepubads.g.doubleclick.net
- static.adsafeprotected.com
- static.doubleclick.net
- stats.g.doubleclick.net
- sync.go.sonobi.com
- s.ytimg.com
- tags.bluekai.com
- tpc.googlesyndication.com
- uipglob.semasio.net
- www.facebook.com
- www.googleadservices.com
- www.google-analytics.com
- www.google.com
- www.google.co.uk
- www.googletagservices.com
- www.youtube.com
- 7senobrlslpagqt8kh6vbxl92ahwu1528731245.nui
- bxszji1mxmxq5qq9g5dtcxaspq6831528730887.nu

The Guardian Content Sources

- assets.guim.co.uk
- i.guim.co.uk
- interactive.guim.co.uk
- ipv4.guim.co.uk
- ipv6.guim.co.uk
- poly_x001c_ll.g
- uim.co.uk
  uploads.guim.co.uk
- api.nextgen.guardianapps.co.uk
- ophan.theguardian.com
- static.theguardian.com
- ad.360yield.com
- adservice.google.com
- adservice.google.co.uk
- apex.go.sonobi.com
- as-sec.casalemedia.com
- beacon.gu-web.net
- beacon.krxd.net
- cdn.adsafeprotected.com
- cdn-gl.imrworldwide.com
- cdn.krxd.net
- connect.facebook.net
- consumer.krxd.net
- dt.adsafeprotected.com
- dual.guim.co.uk
- elb.the-ozone-project.com
- googleads.g.doubleclick.net
- i.ytimg.com
- pagead2.googlesyndication.com
- partner.mediawallahscript.com
- pixel.adsafeprotected.com
- prebid.adnxs.com
- sb.scorecardresearch.com
- secure-au.imrworldwide.com
- secure-dcr.imrworldwide.com
- secure-gl.imrworldwide.com
- securepubads.g.doubleclick.net
- static.adsafeprotected.com
- static.doubleclick.net
- stats.g.doubleclick.net
- sync.go.sonobi.com
- s.ytimg.com
- tags.bluekai.com
- tpc.googlesyndication.com
- uipglob.semasio.net
- www.facebook.com
- www.googleadservices.com
- www.google-analytics.com
- www.google.com
- www.google.co.uk
- www.googletagservices.com
- www.youtube.com
- 7senobrlslpagqt8kh6vbxl92ahwu1528731245.nui
- bxszji1mxmxq5qq9g5dtcxaspq6831528730887.nu

The Telegraph Content Sources

static.telegraph.co.uk
assets.adobedtm.com
dpm.demdex.net
img.youtube.com
sb.scorecardresearch.com
zm232.com
secure.s.telegraph.co.uk
c.amazon-adsystem.com
www.googletagservices.com
tmg.demdex.net
securepubads.g.doubleclick.net
connect.facebook.net
www.google-analytics.com
adservice.google.com
adservice.google.co.uk
www.facebook.com
js-sec.indexww.com
bat.bing.com
telegraph.sdk.beemray.com
hm732.com
amplify.outbrain.com
ib.adnxs.com
telegraphmedia.bootstrap.fyre.co
amplifypixel.outbrain.com
aax.amazon-adsystem.com
cdn.taboola.com
tr.outbrain.com
telegraphmediagroupl.tt.omtrdc.net
telegraph.grapeshot.co.uk
ict.in_x001c_nity-tracking.net
cdn.petametrics.com
collector-2794.tvsquared.com
acdn.adnxs.com
advertising.oriel.io
static.criteo.net
opentag-stats.qubit.com

v7.beemray.com
as-sec.casalemedia.com
secure.adnxs.com
trc.taboola.com
query.petametrics.com
bidder.criteo.com
router.infolinks.com
load77.exelator.com
pr-bh.ybp.yahoo.com
aax-eu.amazon-adsystem.com
sync.search.spotxchange.com
beacon.walmart.com
s.amazon-adsystem.com
m.adnxs.com
pagead2.googlesyndication.com
tpc.googlesyndication.com
pixel.advertising.com
www.google.com
secure-assets.rubiconproject.com
cdn.ampproject.org
ams1-ib.adnxs.com
plugin.mediavoice.com
eus.rubiconproject.com
native.sharethrough.com
meraxes-cdn.polarmobile.com
ads.pubmatic.com
fw.adsafeprotected.com
cdn.adnxs.com
us-u.openx.net
pixel.adsafeprotected.com
gb-gmtdmp.mookie1.com
z.moatads.com
cdn.mookie1.com
static-tagr.gd1.mookie1.com
d.lemonpi.io
match.adsrvr.org

dt.adsafeprotected.com
px.moatads.com
eu-u.openx.net
b.sharethrough.com
googleads.g.doubleclick.net
token.rubiconproject.com
ad.doubleclick.net
cdn.mediavoice.com
image6.pubmatic.com
odr.mookie1.com
pixel-eu.rubiconproject.com
static.adsafeprotected.com
pubads.g.doubleclick.net
s0.2mdn.net
pentos-cdn.polarmobile.com
googleads4.g.doubleclick.net
pubmatic-match.dotomi.com
c1.adform.net
pubmatic2waycm-atl.netmng.com
cdnjs.cloud_x001d_are.com
image4.pubmatic.com
simage4.pubmatic.com
p.skimresources.com
static.chartbeat.com
s3-eu-west-1.amazonaws.com
ping.chartbeat.net
secure.widget.cloud.opta.net
www.google.co.uk
ssl.google-analytics.com
ade.googlesyndication.com
pixel.tapad.com
pub.pxl.ace.advertising.com
simage2.pubmatic.com
image2.pubmatic.com

cm.ctnsnet.com
connexity.net
magnetic.t.domdex.com
pixel.rubiconproject.com
dis.criteo.com
r.turn.com
t.wayfair.com
p.adsymptotic.com
delivery.h.switchadhub.com
choices.truste.com
eur-ukp.adsrvr.org
ad.atdmt.com
a.ctnsnet.com
cdn.ctnsnet.com
cm.g.doubleclick.net
choices.trustarc.com
code.createjs.com
collector.cint.com
dsum-sec.casalemedia.com
cm.adform.net
ad.360yield.com
eb2.3lift.com
jadserve.postrelease.com
rtb-csync.smartadserver.com
sync.teads.tv
dh.serving-sys.com
s.thebrighttag.com
adserver-us.adtech.advertising.com
www.national-lottery.co.uk
d3c3cq33003psk.cloudfront.net
telegraphmediagroup.d3.sc.omtrdc.net
performance-data.gcpdata.telegraph.co.uk

BBC Content Sources

- nav.files.bbci.co.uk
- static.bbc.co.uk
- homepage.files.bbci.co.uk
- ichef.bbci.co.uk
- g.bbc.co.uk
- mybbc.files.bbci.co.uk
- ssl.bbc.co.uk
- static.bbci.co.uk
- search.files.bbci.co.uk
- mvt.api.bbc.com
- sa.bbc.co.uk
- idcta.api.bbc.co.uk
- edigitalsurvey.com

The Guardian Content Sources

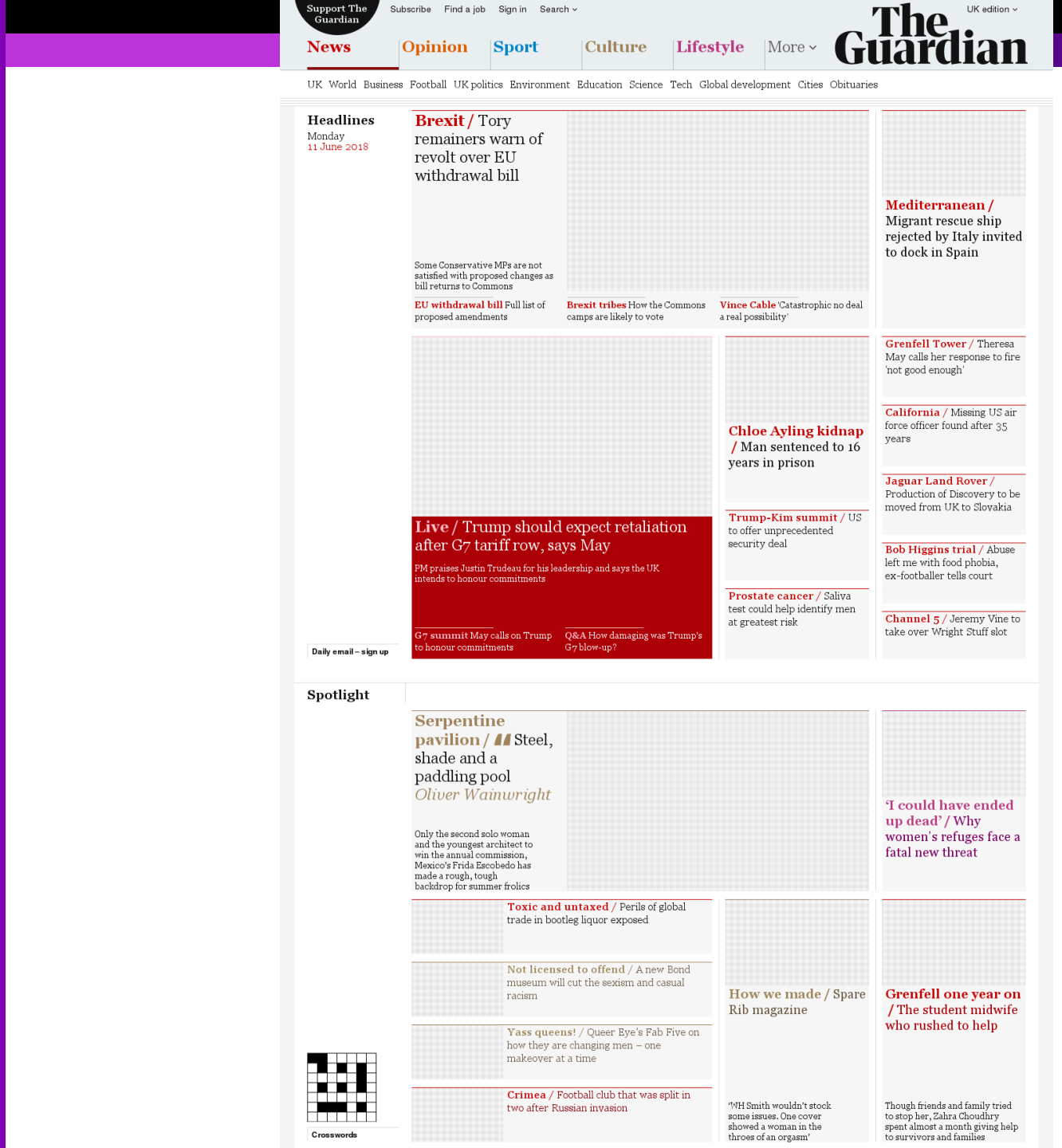
- assets.guim.co.uk
- i.guim.co.uk
- interactive.guim.co.uk
- ipv4.guim.co.uk
- ipv6.guim.co.uk
- poly_x001c_ll.g
- uim.co.uk
  uploads.guim.co.uk
- api.nextgen.guardianapps.co.uk
- ophan.theguardian.com
- static.theguardian.com
- ad.360yield.com
- adservice.google.com
- adservice.google.co.uk
- apex.go.sonobi.com
- as-sec.casalemedia.com
- beacon.gu-web.net
- beacon.krxd.net
- cdn.adsafeprotected.com
- cdn-gl.imrworldwide.com
- cdn.krxd.net
- connect.facebook.net
- consumer.krxd.net
- dt.adsafeprotected.com
- dual.guim.co.uk
- elb.the-ozone-project.com
- googleads.g.doubleclick.net
- i.ytimg.com
- pagead2.googlesyndication.com
- partner.mediawallahscript.com
- pixel.adsafeprotected.com
- prebid.adnxs.com
- sb.scorecardresearch.com
- secure-au.imrworldwide.com
- secure-dcr.imrworldwide.com
- secure-gl.imrworldwide.com
- securepubads.g.doubleclick.net
- static.adsafeprotected.com
- static.doubleclick.net
- stats.g.doubleclick.net
- sync.go.sonobi.com
- s.ytimg.com
- tags.bluekai.com
- tpc.googlesyndication.com
- uipglob.semasio.net
- www.facebook.com
- www.googleadservices.com
- www.google-analytics.com
- www.google.com
- www.google.co.uk
- www.googletagservices.com
- www.youtube.com
- 7senobrlslpagqt8kh6vbxl92ahwu1528731245.nui
- bxszji1mxmxq5qq9g5dtcxaspq6831528730887.nu

The Guardian as you would normally see it

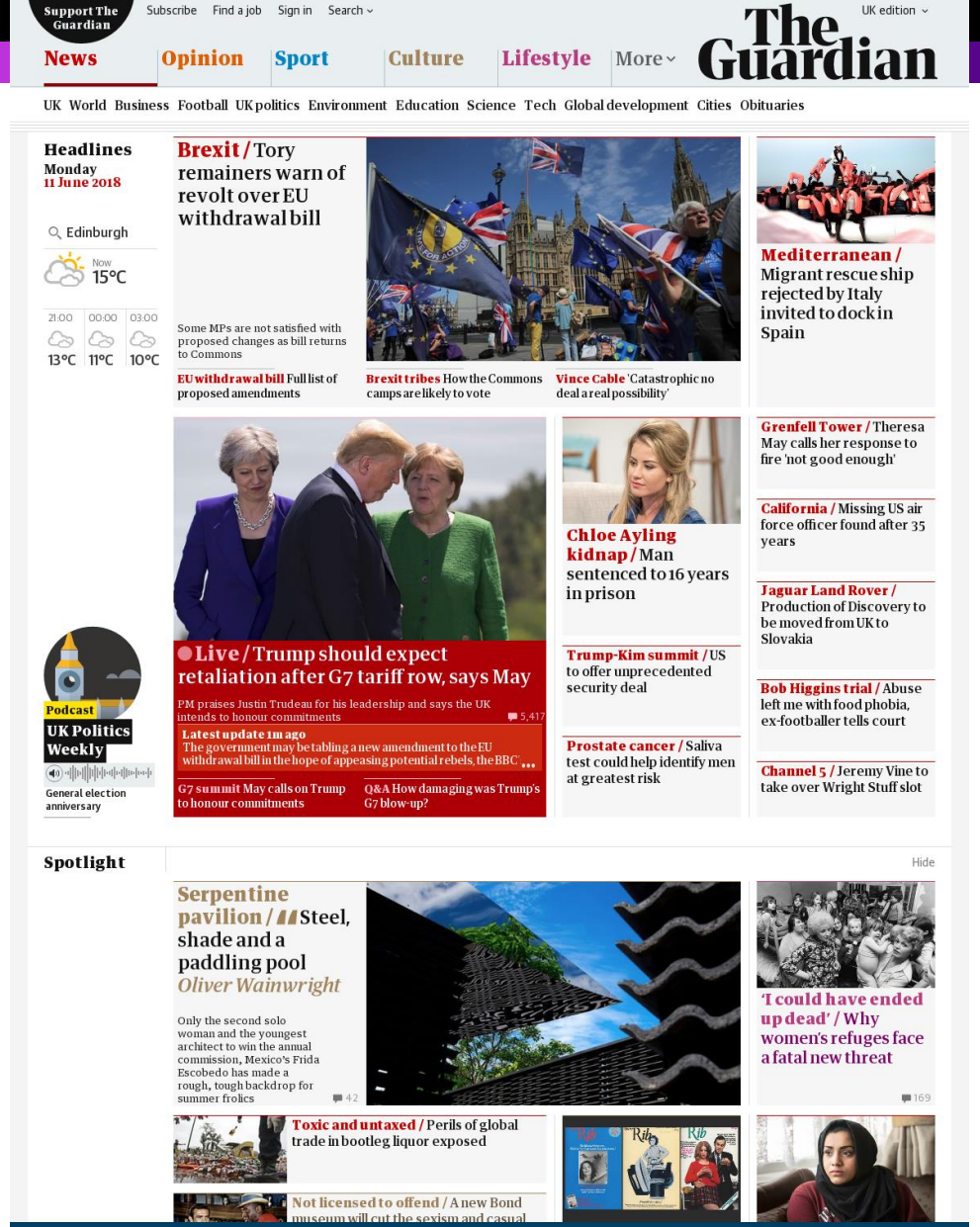
✓ theguardian.co.uk

X Guardian owned

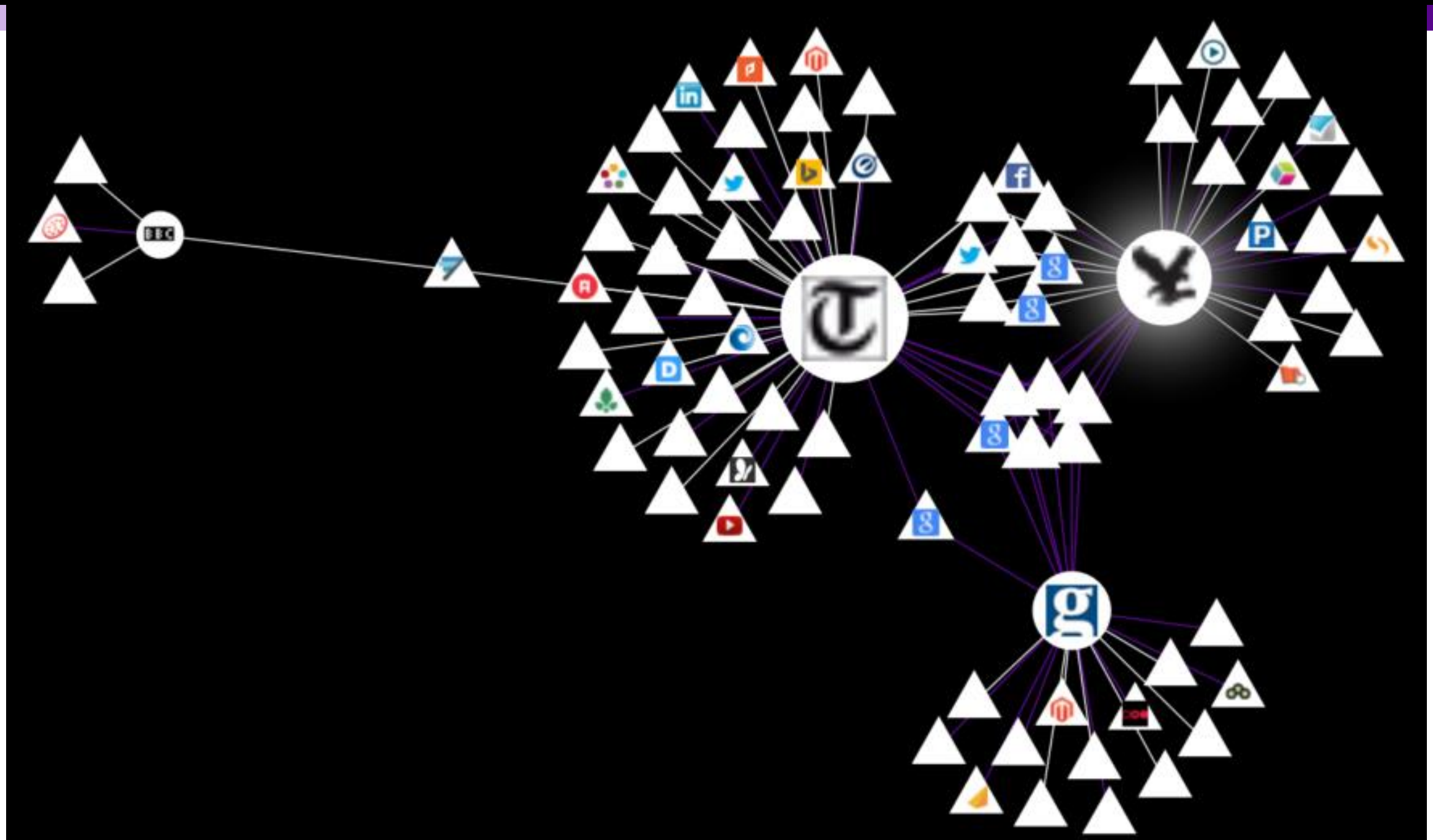
X Trackers



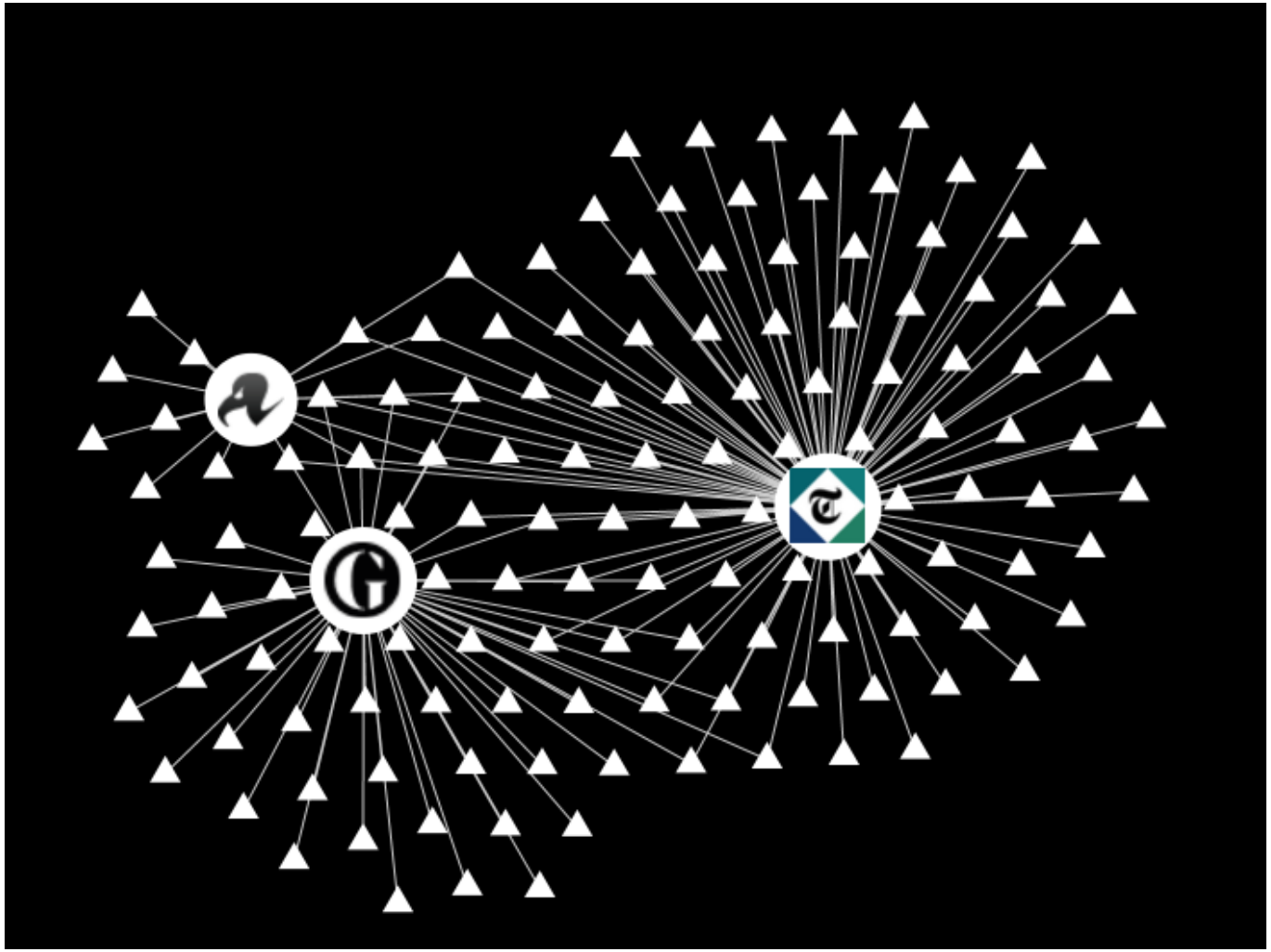
- ✓ theguardian.co.uk
- ✓ Guardian owned
- X Trackers



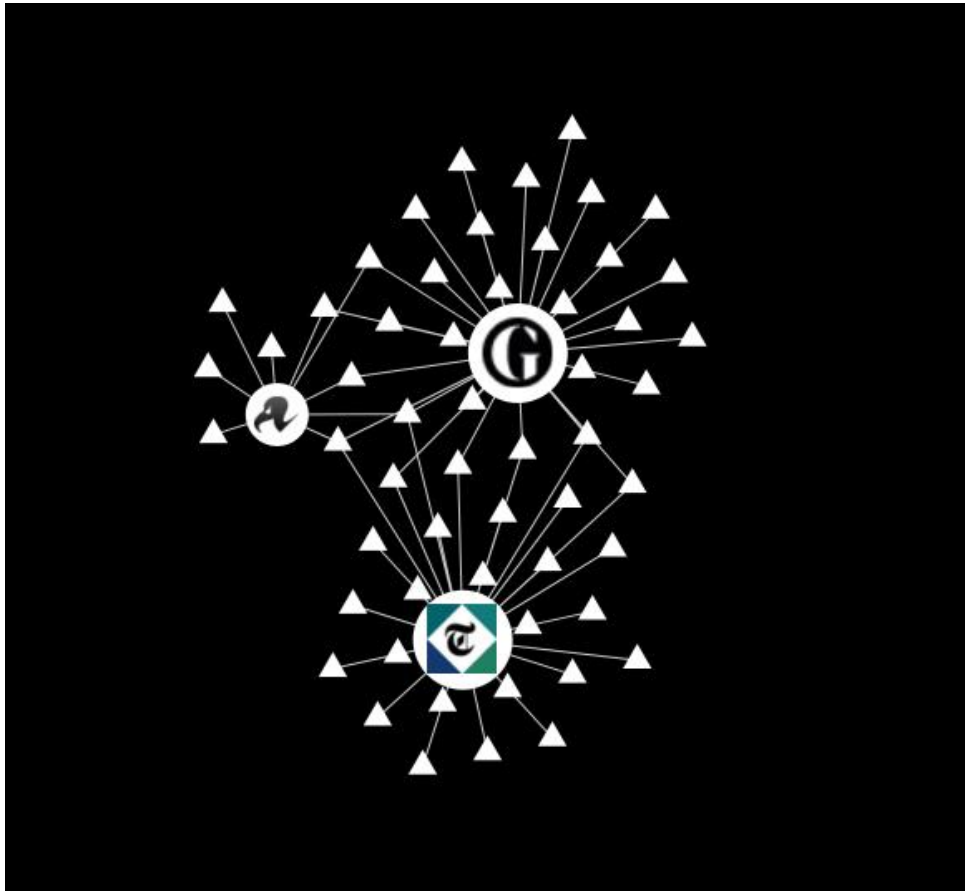
- ✓ theguardian.co.uk
- ✓ Guardian owned
- ✓ Trackers


BBC

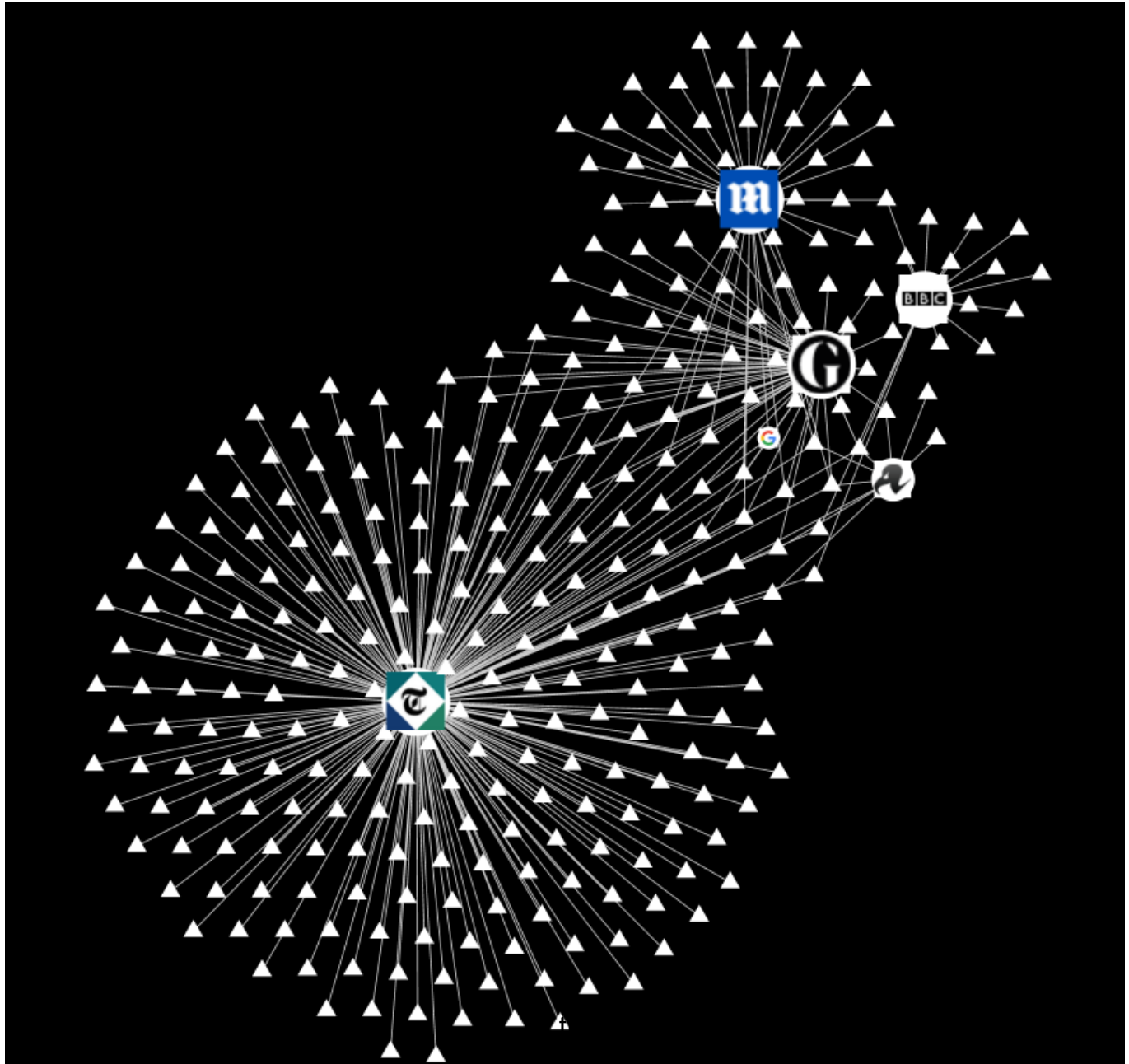
Normal



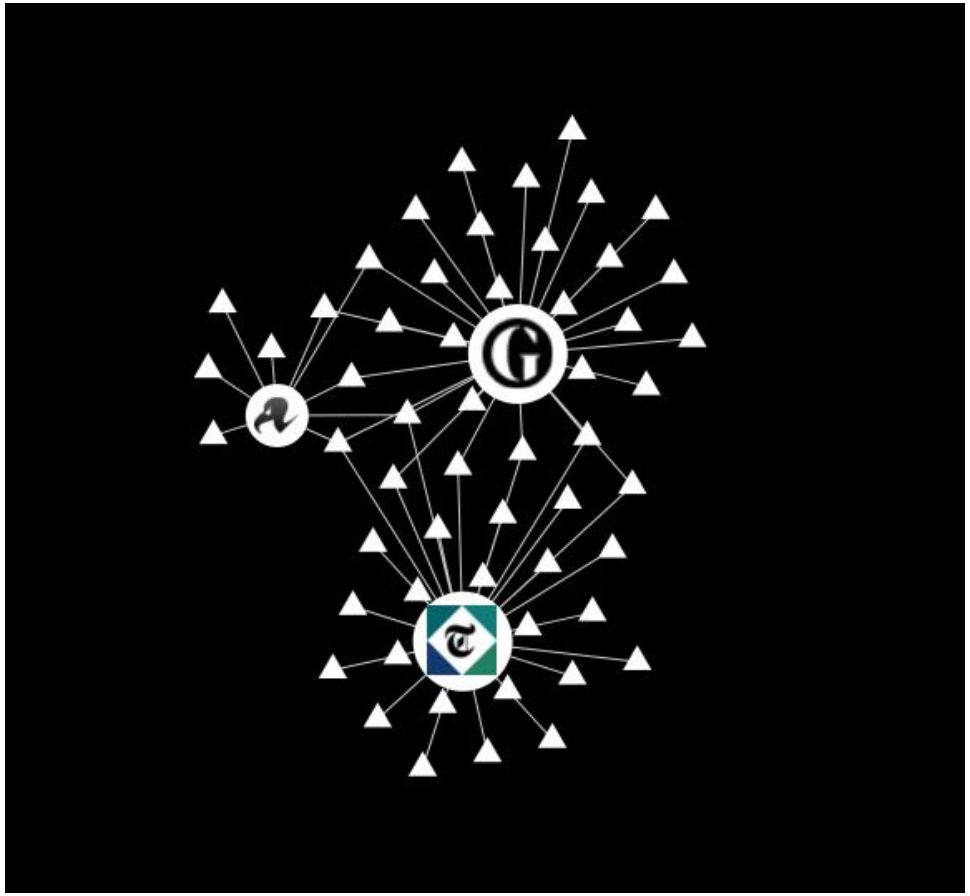
Adblock Plus



Normal (hours later)



Adblock Plus



Cute Dogs!

Cute dog sleeping on a park bench.

Playing with a ball all day is hard work. But that is no reason to release the ball.

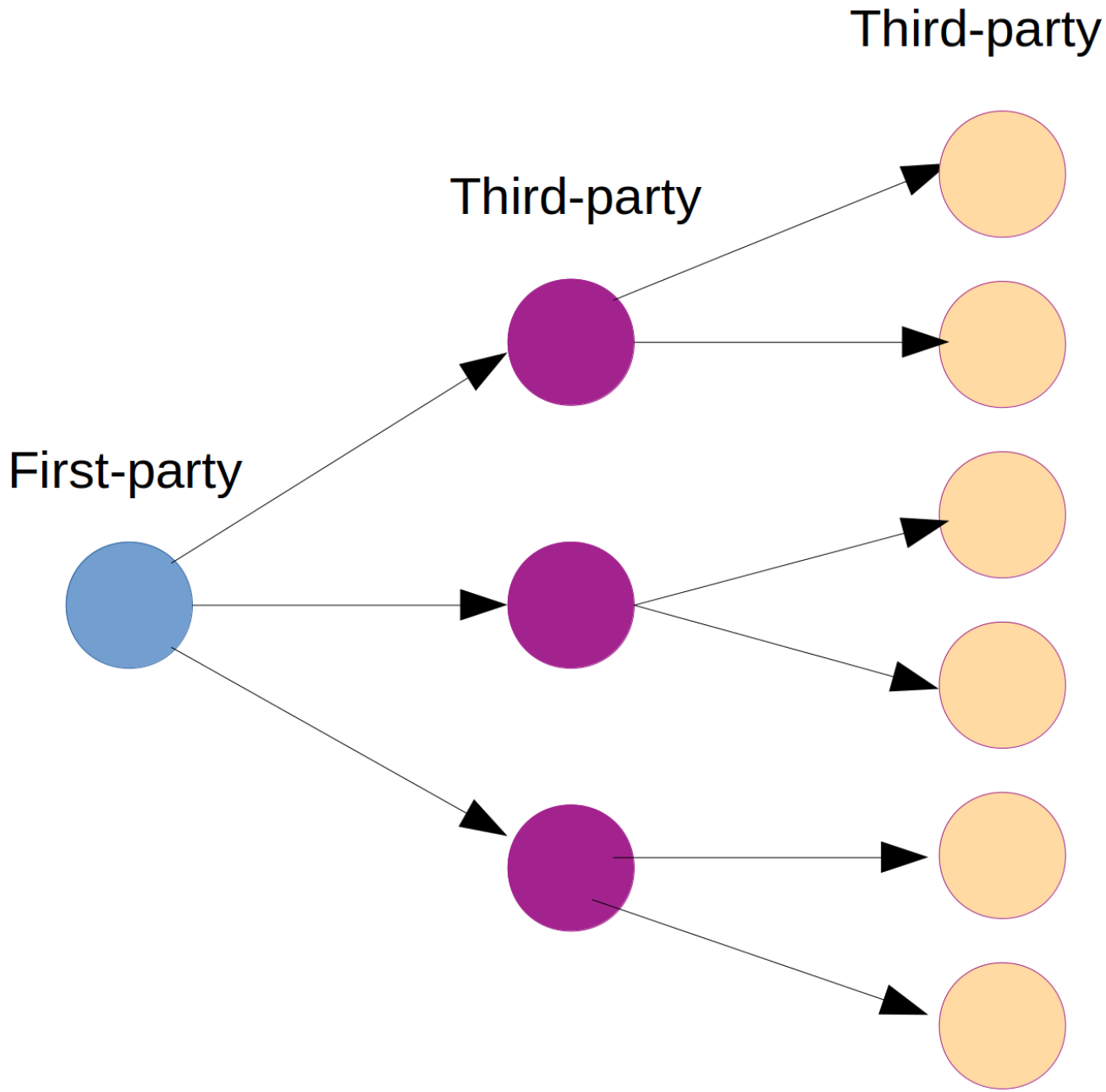
Think-pair-share

- What does an advertising blocker do in terms of the page loading sequence we just saw?
- How could DNS be used to do advertising blocking?
- Advanced:
  - How do advertising companies like Facebook and Google verify ads on their platforms are “safe”?

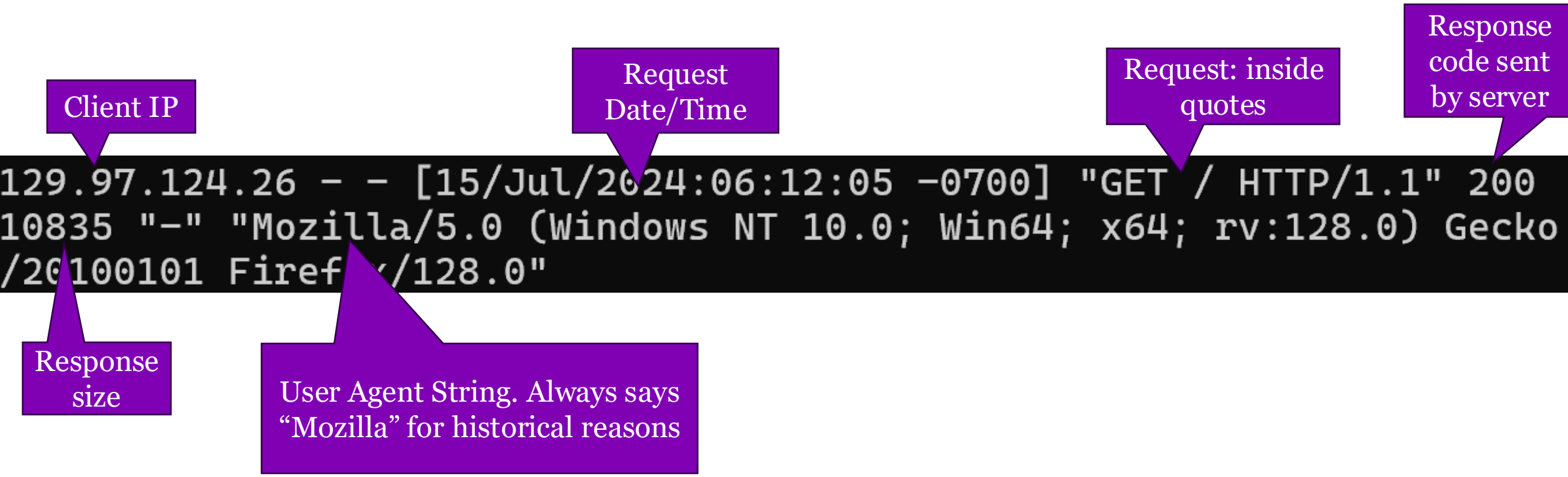
WEB PAGE BUILDING: WEB REQUESTS

Each piece of content can add new content to the page.

Every content request sends data to the web server.



HTTP Request: Every connection sends information



HTTP Request: Every connection sends information

- Firefox on Windows

```
129.97.124.26 - - [15/Jul/2024:06:12:05 -0700] "GET / HTTP/1.1" 200
10835 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:128.0) Gecko
/20100101 Firefox/128.0"
```

- Firefox on Android

```
129.97.124.32 - - [15/Jul/2024:06:26:32 -0700] "GET / HTTP/1.1" 200
10835 "-" "Mozilla/5.0 (Android 14; Mobile; rv:128.0) Gecko/128.0 Fi
refox/128.0"
```

- Edge on Windows

```
129.97.124.26 - - [15/Jul/2024:06:23:56 -0700] "GET / HTTP/1.1" 200
10835 "-" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537
.36 (KHTML, like Gecko) Chrome/126.0.0.0 Safari/537.36 Edg/126.0.0.0
"
```

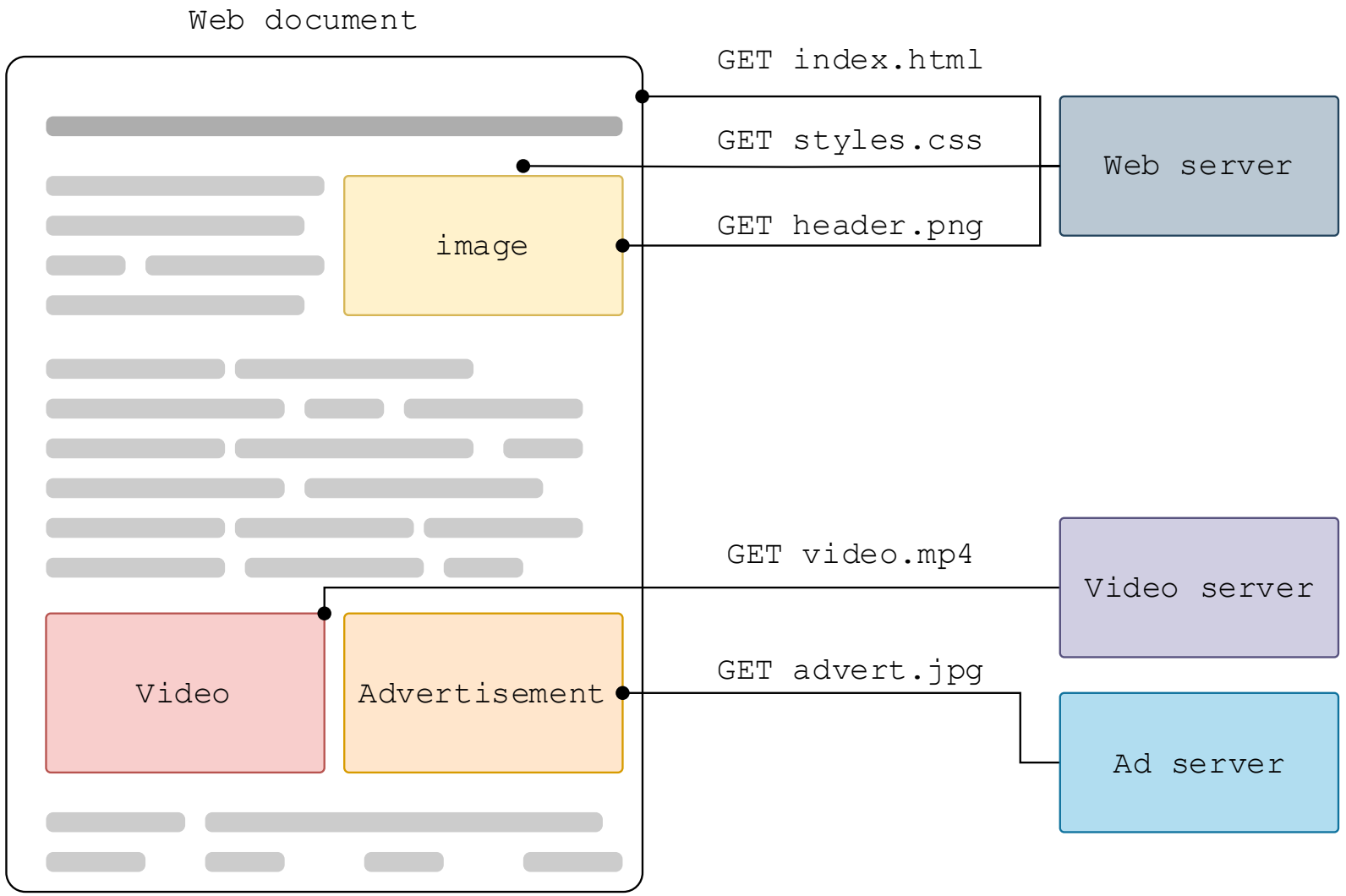
Log of an attack attempt

My website has no PHP installed.
This is an attack attempt.

```
119.188.247.39 - - [15/Jul/2024:06:52:49 -0700] "GET /login.php?s=Admin
/login HTTP/1.1" 404 3856 "-" "python-requests/2.31.0"
223.109.252.233 - - [15/Jul/2024:06:55:40 -0700] "GET / HTTP/1.1" 200 3
5214 "-" "Sogou web spider/4.0(+http://www.sogou.com/docs/help/webmaste
rs.htm#07)"
```

This line is an automatic script downloading content off my website.

Every connection sends information. Every connection constructed with a scripting language (i.e. Javascript) can dynamically add GET/POST data to the request that it gets locally from the DOM.

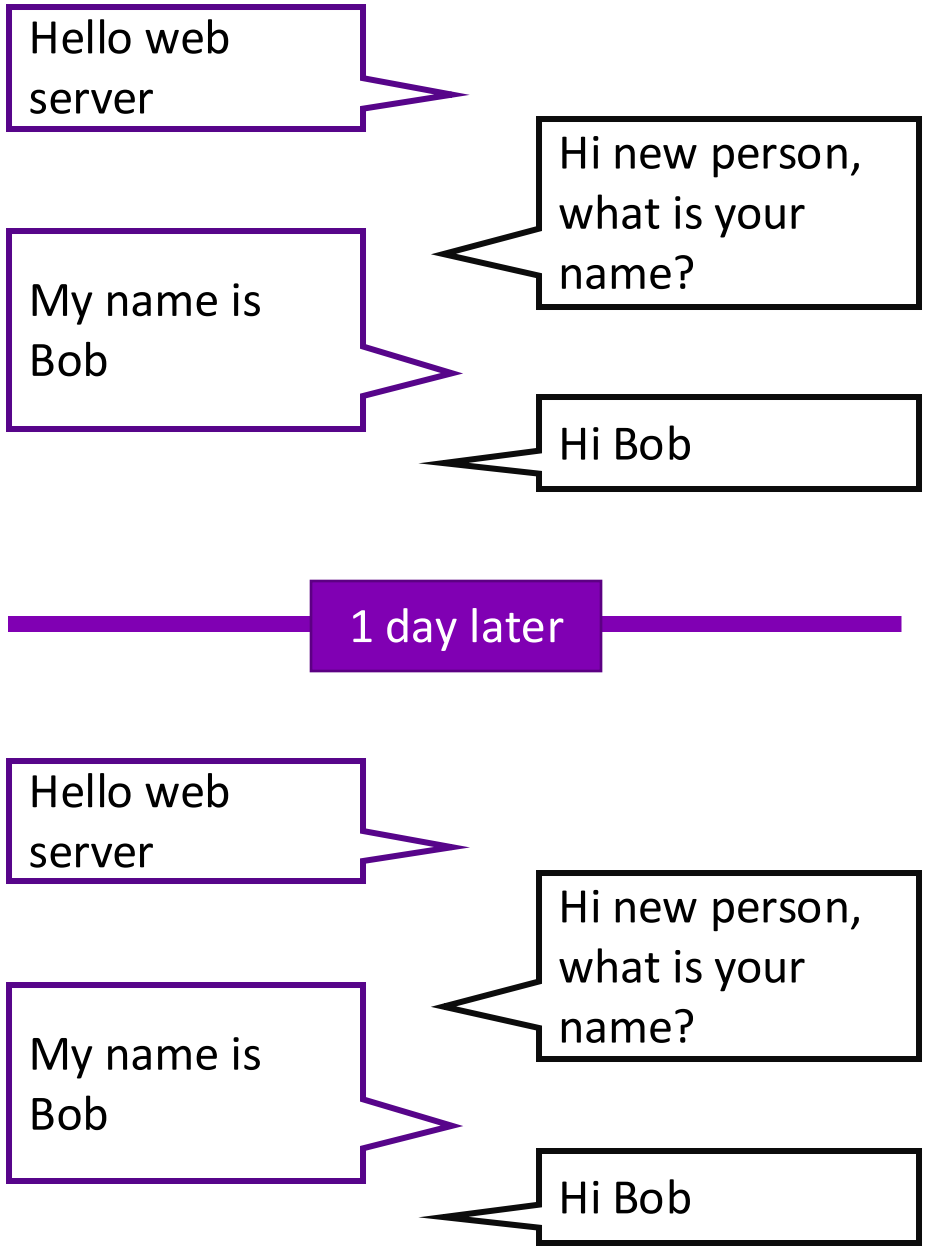


https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview

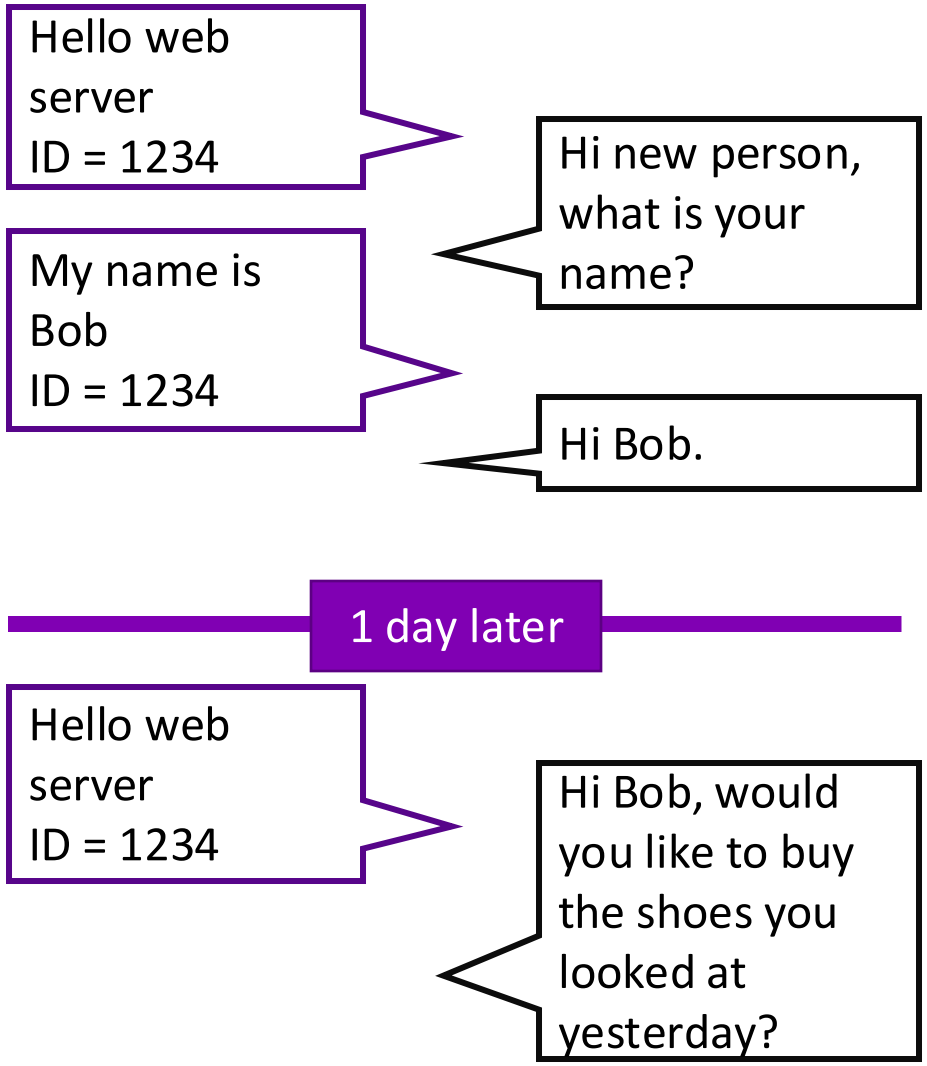
DESIGNING THE WEB COOKIE

Heavily based on Lou Montulli's "The reasoning behind Web Cookies"

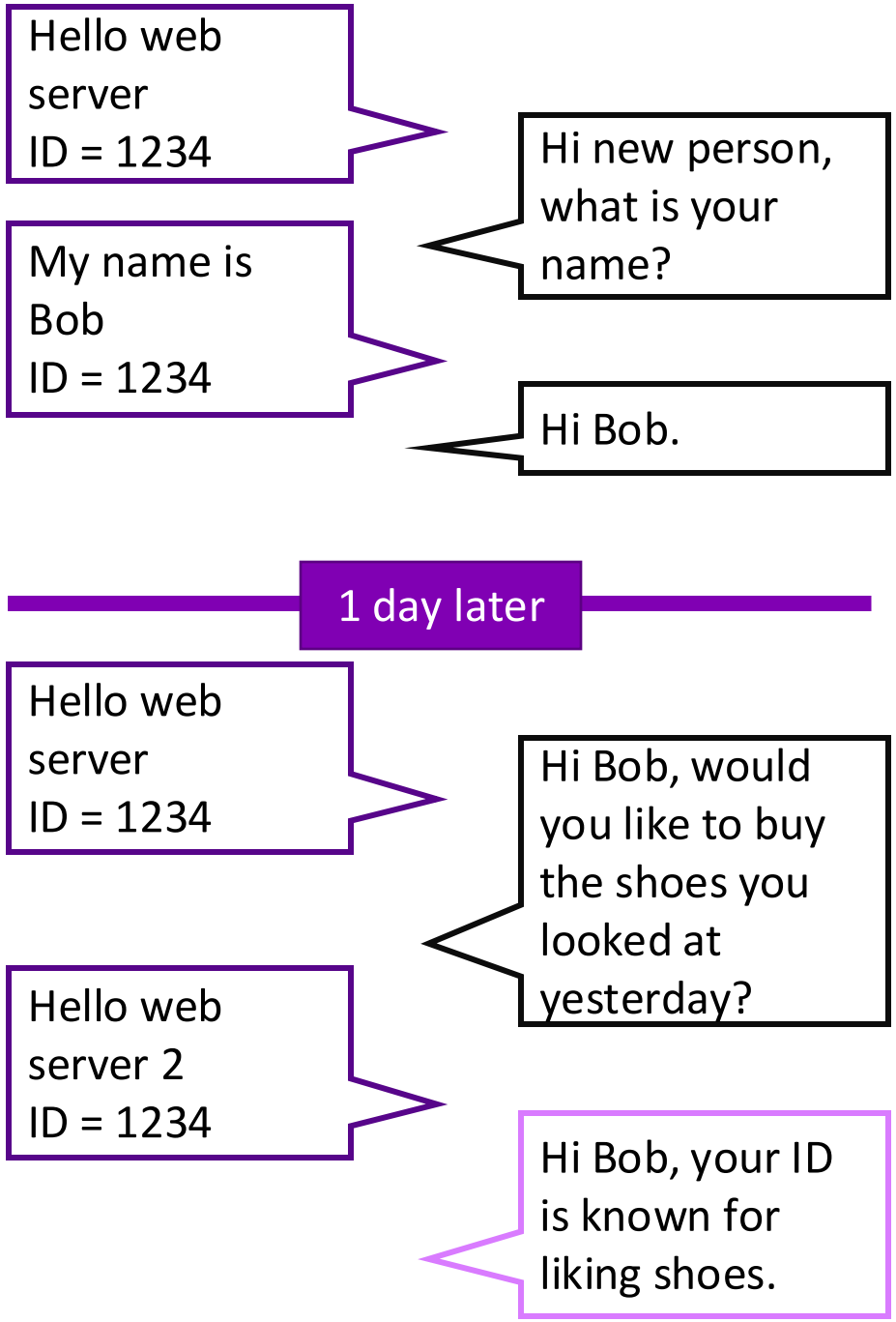
The year is 1994 and there is a problem... the internet has no ability to remember a person between page reloads.



There is an obvious easy solution… Give each browser a unique identifier.

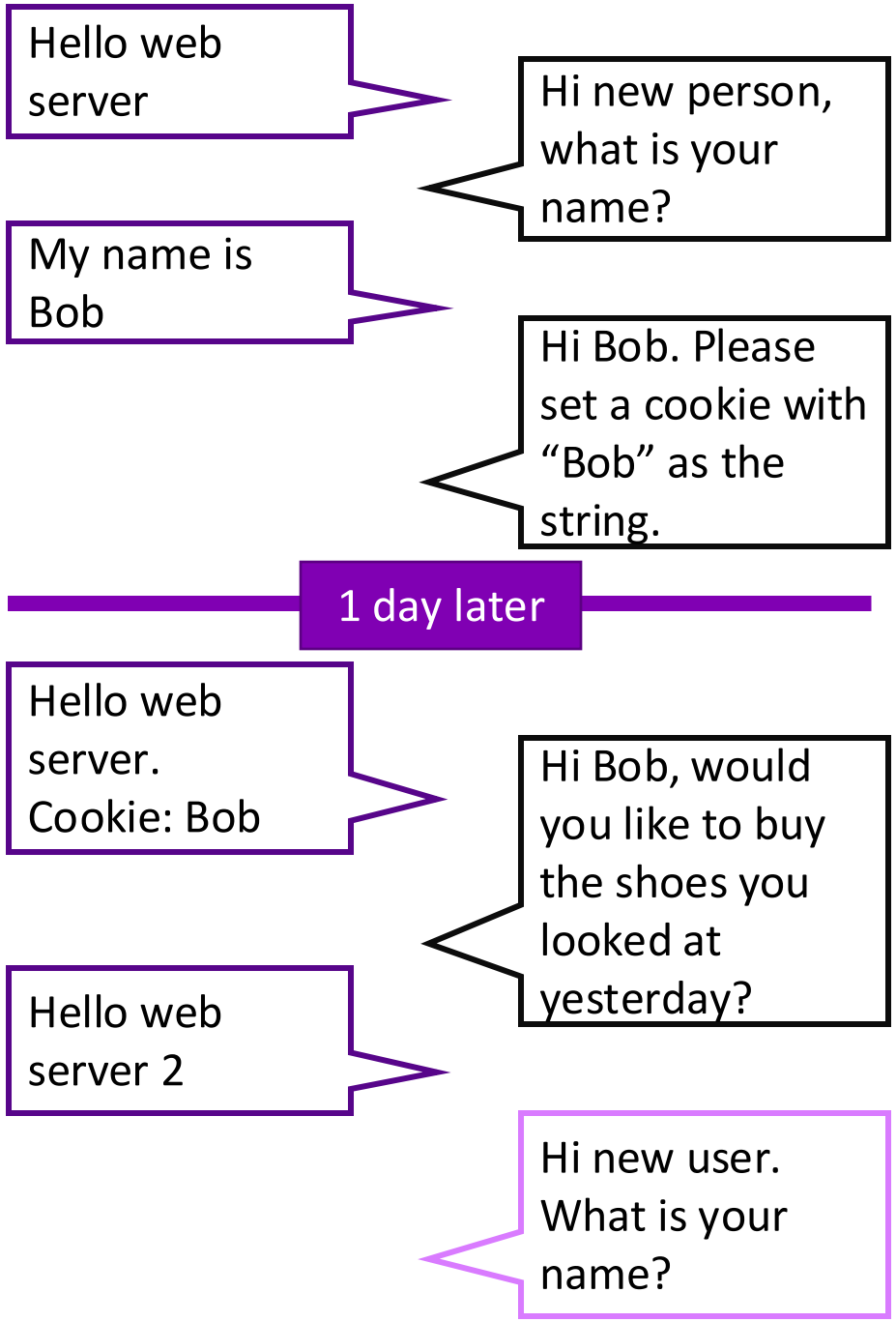


The problem with the obvious solution is privacy. Tracking would be possible with no visibility or control.



Instead Netscape implemented cookies.

Small text strings the server can ask the browser to remember and give back later.



QUESTIONS