Ollscoil na Gaillimhe
University of Galway

CT 420
Real-Time Systems

Emerging Protocols-I

Dr. Jawad Manzoor
Assistant Professor
School of Computer Science

University *of* Galway.ie

Contents

Ollscoil na Gaillimhe
University of Galway

- Web QoS
- Achieving Speed
- Evolution of the web
- Improvements in HTTP

QoS Attribute of Web Services

Ollscoil na Gaillimhe
University of Galway

QoS Attribute	Definition
Accuracy	Error rate produced by the service
Accessibility	Degree the service is capable of serving a Web service request
Capacity	Limit of concurrent requests for guaranteed performance
Response Time	Time to complete a Web service request (from a client perspective)
Throughput	Number of Web service requests served at a given time period
Availability	The probability that the service can respond to the consumer requests
MTTR	Meantime to repair
Interoperability	The ease with which a consumer application or agent interoperates with a service
Robustness	The degree to which a service can function correctly in the presence of invalid, incomplete or conflicting inputs
Authentication	A measure of how the service authenticate principals who can access service and data
Confidentiality	A measure of how the service threat the data, so that only authorized principals can access or modify the data

Website speed

Ollscoil na Gaillimhe
University of Galway

- Website speed, or website performance, refers to how quickly a browser is able to load fully functional webpages from a given site.

- Why is site speed important?
  - User experience
  - SEO
  - Bounce rate

Web Performance Translates to Revenue

Ollscoil na Gaillimhe
University of Galway

- Well-publicized studies from Google, Microsoft, and Amazon all show that web performance translates directly to revenue
  - e.g., a 2,000 ms delay on Bing search pages decreased per-user revenue by 4.3%!

- Another study of over 160 organizations determined that:
  - an extra one-second delay in page load times led to 7% loss in conversions, 11% fewer page views, and a 16% decrease in customer satisfaction!

Web Performance Translates to Revenue

Ollscoil na Gaillimhe
University of Galway

- Mobify found that decreasing their homepage's load time by 100 milliseconds resulted in a 1.11% uptick in session-based conversion
- Retailer AutoAnything experienced a 12-13% increase in sales after cutting page load time in half
- Walmart discovered that improving page load time by one second increased conversions by 2%

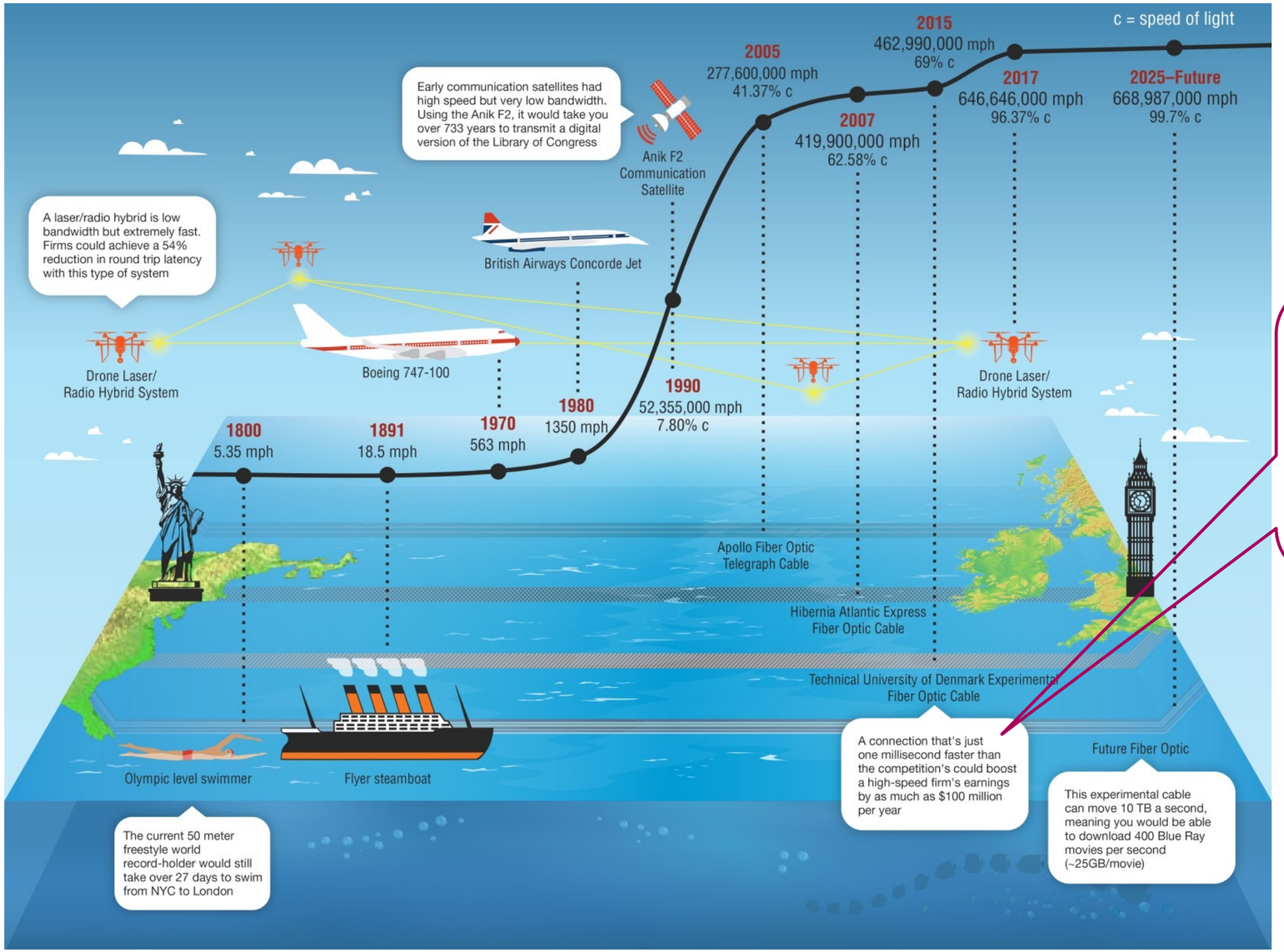
- *Faster sites yield more page views, higher engagement, and higher conversion rates.*

How to achieve speed?

Ollscoil na Gaillimhe
University of Galway

- There are rules to the universe.
- Information has a speed limit:
  - namely, c, the speed of light
  - light travels about 300,000,000 meters per second.

Improving the communication link



A connection that's just one millisecond faster than the competition's could boost a high-speed firm's earnings by as much as $100 million per year

https://techcrunch.com/2015/11/08/data-the-speed-of-light-and-you/

Improving the communication link

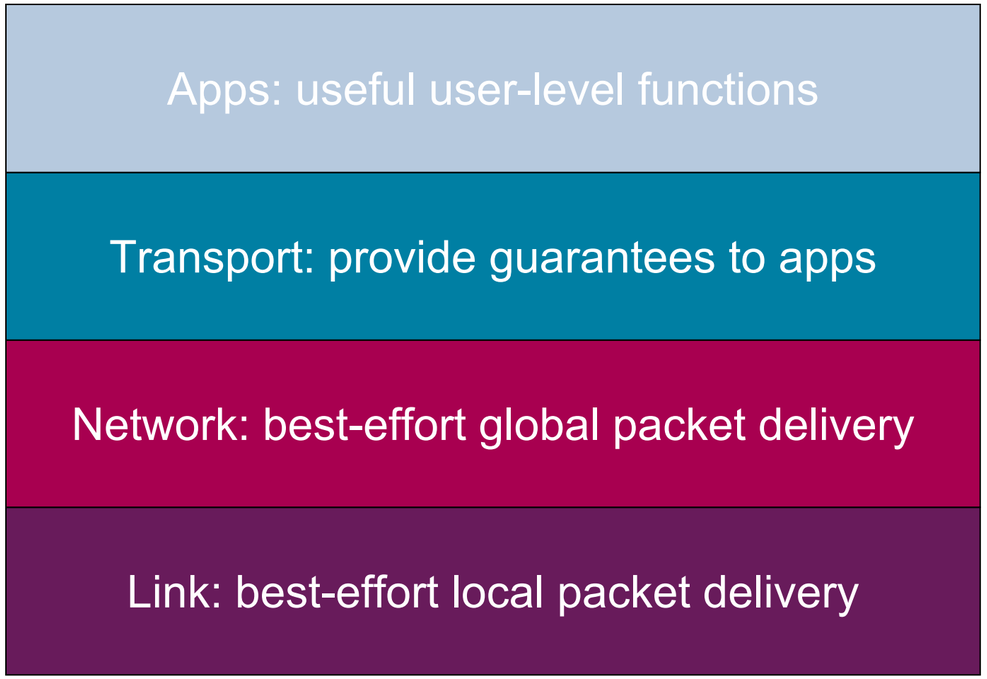
- Google’s "Dunant" trans-Atlantic cable has the capacity to deliver a massive 250 terabits per second. (2021)
- The "Grace Hopper" cable has total of 352Tbps system capacity. (2022)
- The newly completed transatlantic cable called Amitié and funded by Microsoft, Meta and others, can carry 400 Tbps. (2023)

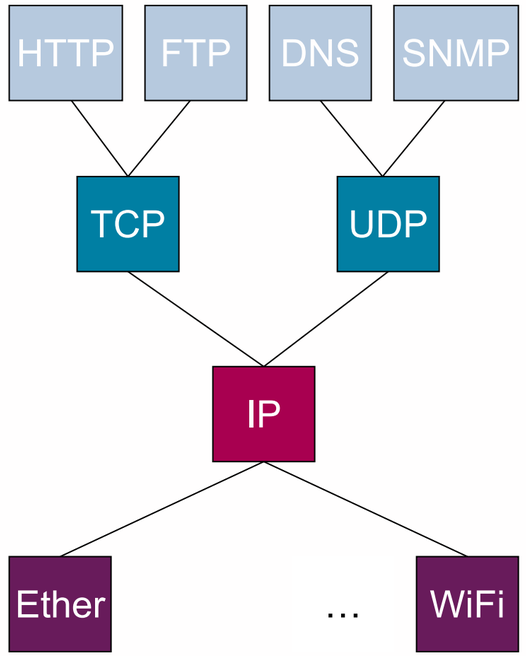
An optical fibre about the thickness of a human hair can now carry the equivalent of more than 10 million fast home internet connections running at full capacity.

A team of Japanese, Australian, Dutch, and Italian researchers has set a new speed record for an industry standard optical fibre, achieving 1.7 Petabits over a 67km length of fibre.

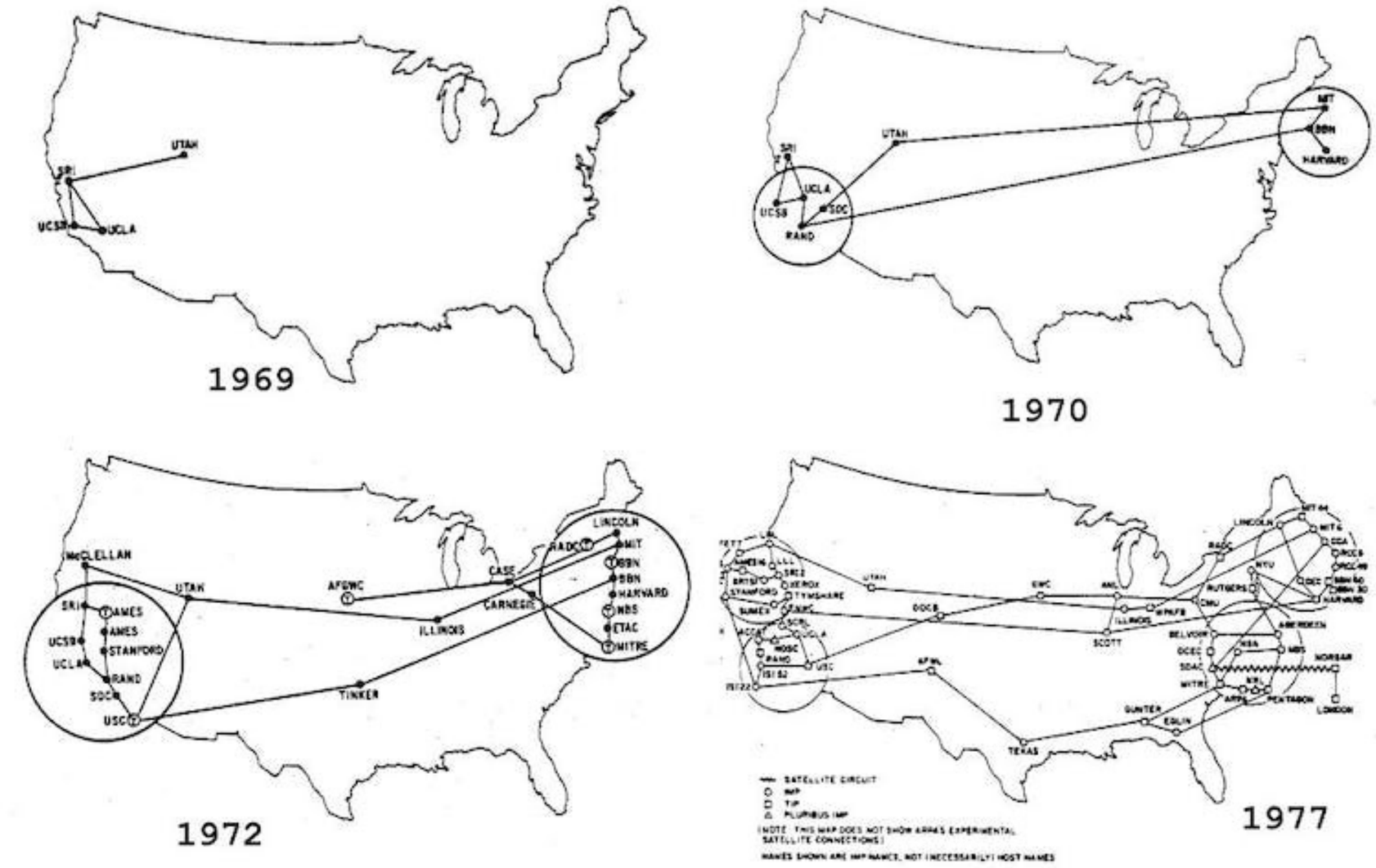
https://www.sciencedaily.com/releases/2023/05/230530125447.htm

Improving the protocols





History of the Internet



ARPANET 1969-1977. Wikipedia

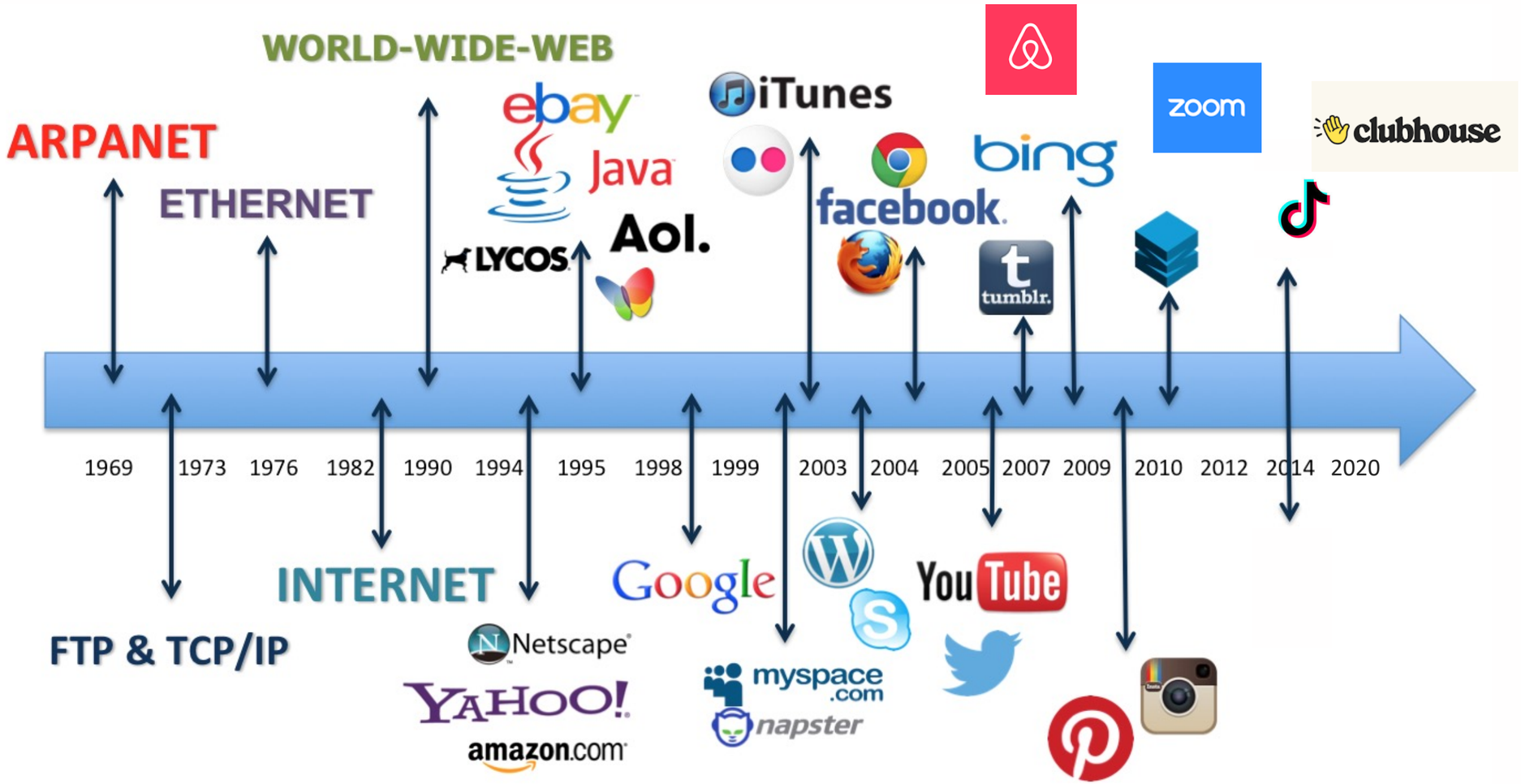
History of the Internet

Ollscoil na Gaillimhe
University of Galway

- 1965 - Lawrence Roberts (MIT) & Thomas Marill create first Wide-Area Network connection via long distant dial-up between computers in Massachusetts and California.
- 1967 - ARPAnet design begins
- 1973 - TCP/IP protocol development begins, headed by Vint Cerf (Stanford) and Robert Kahn (DARPA).
- 1981 - Internet protocol version 4, or IPv4, was officially defined in RFC 791.
- 1989 - Tim Berners-Lee Creates the World Wide Web (WWW)

History of the Internet

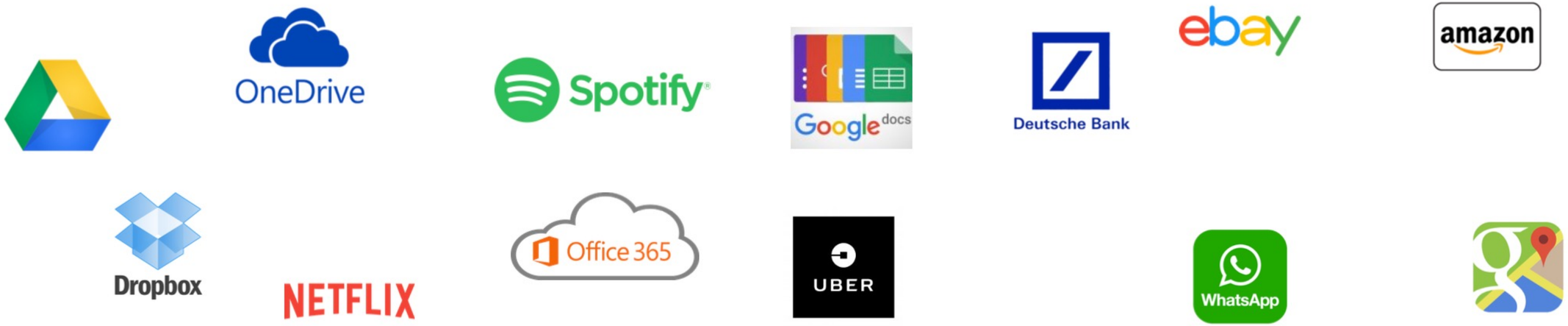
Ollscoil na Gaillimhe
University of Galway



Evolution of the Web

Ollscoil na Gaillimhe
University of Galway

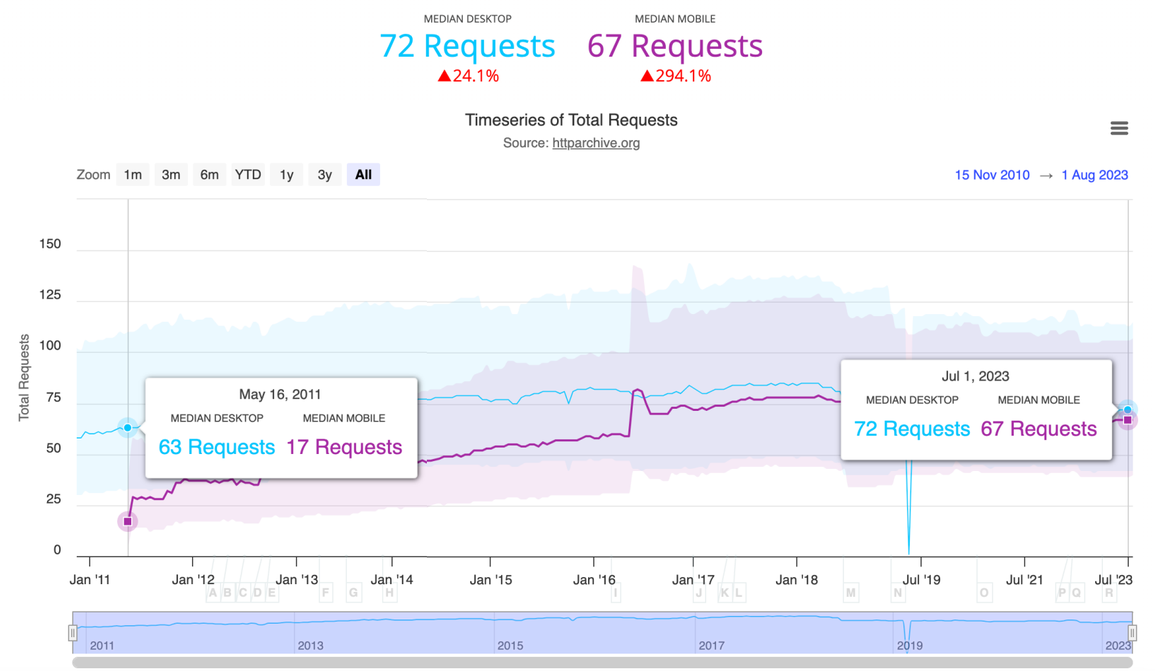
- There is a paradigm shift in the way the Internet is accessed
  - Transformation from native applications to web applications
- Traditional applications are getting migrated to the cloud
- Web applications and services are getting extremely complex
- QoS is web applications is very important

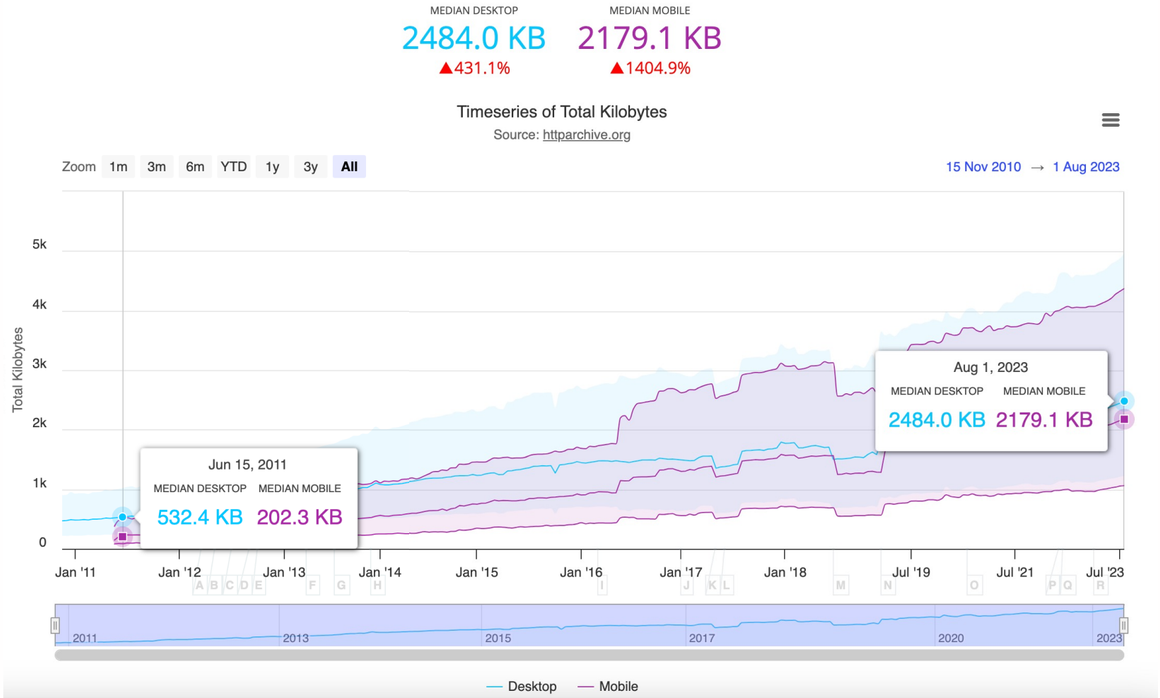


Evolution of the Web

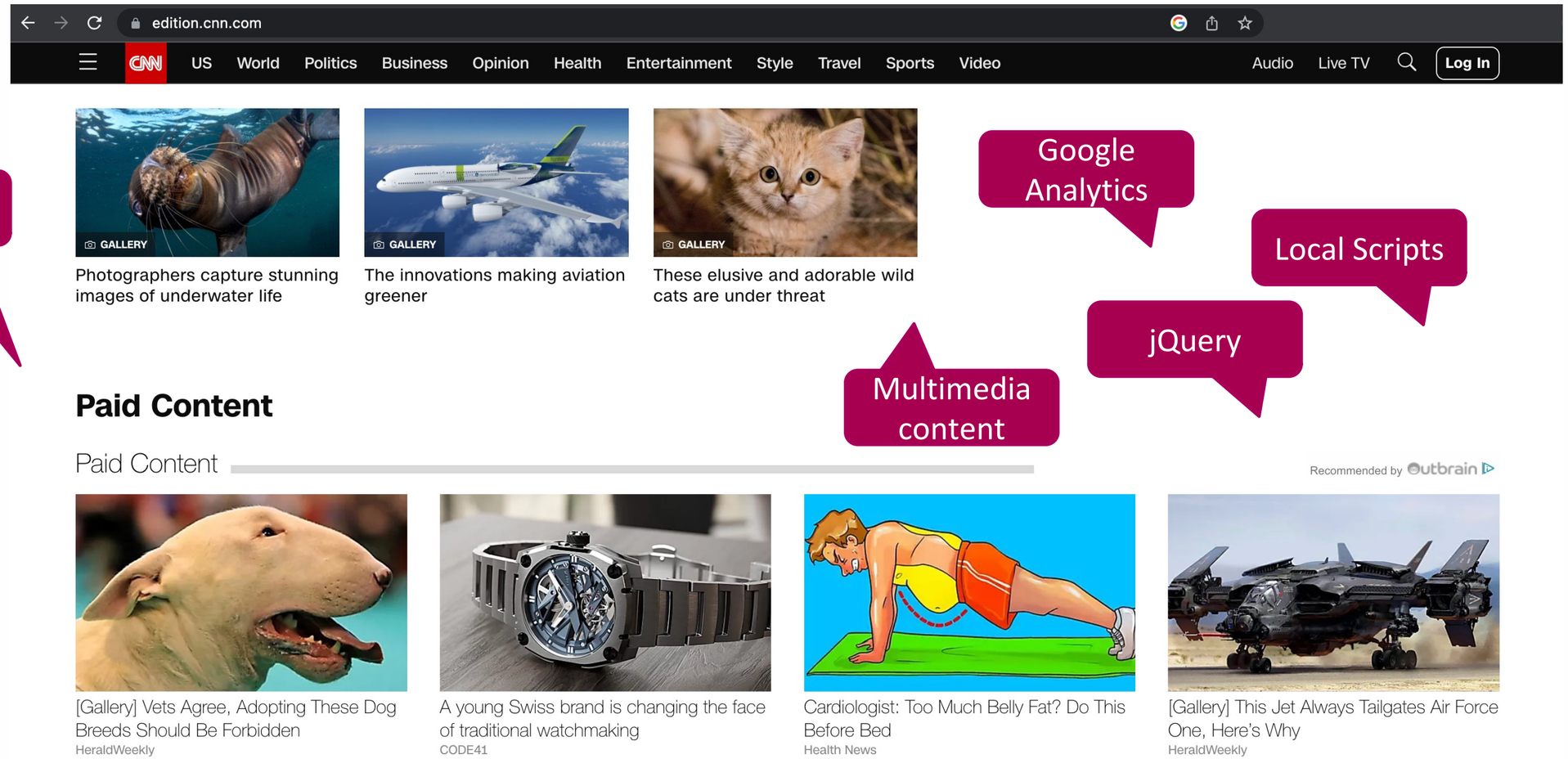
Ollscoil na Gaillimhe
University of Galway

- An average web page size and the number of requests to load a web page has increased manyfold.

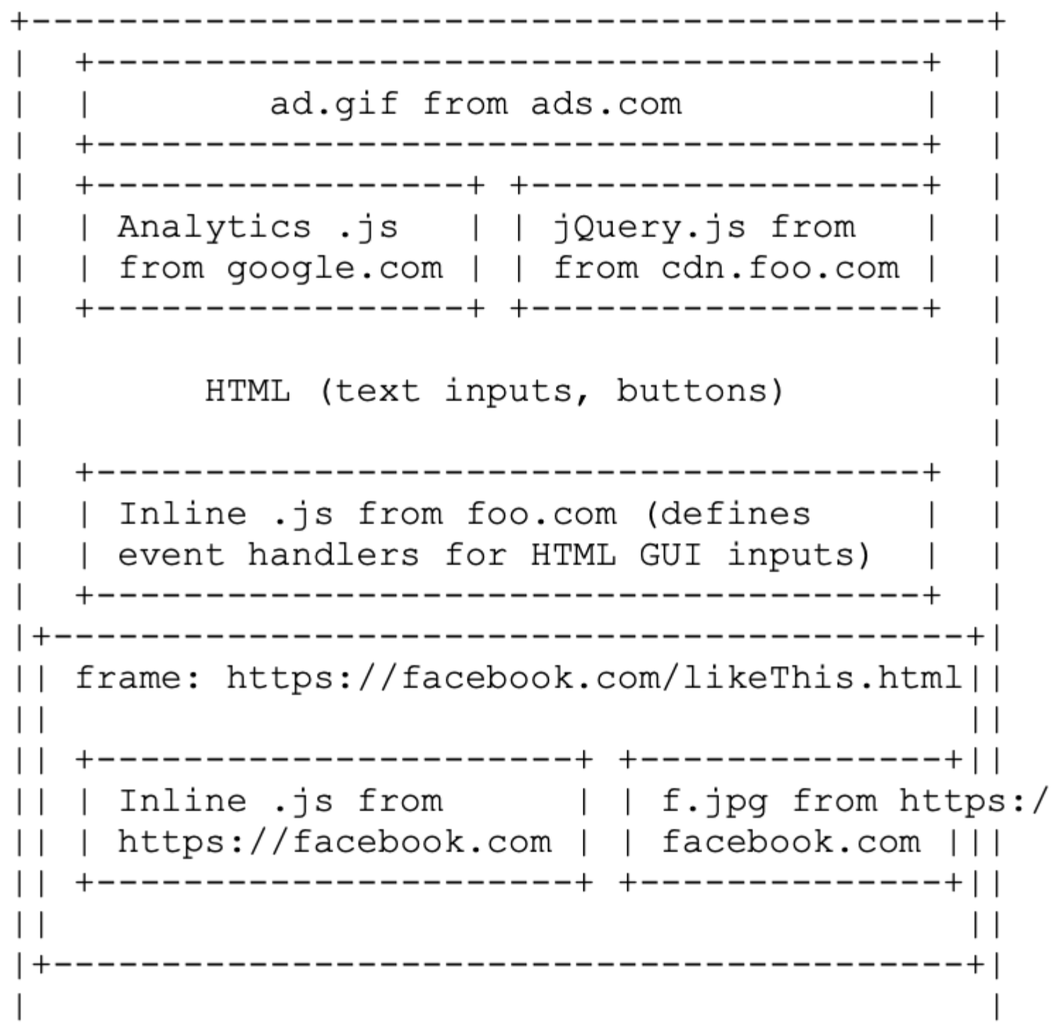




Anatomy of a Modern Web App

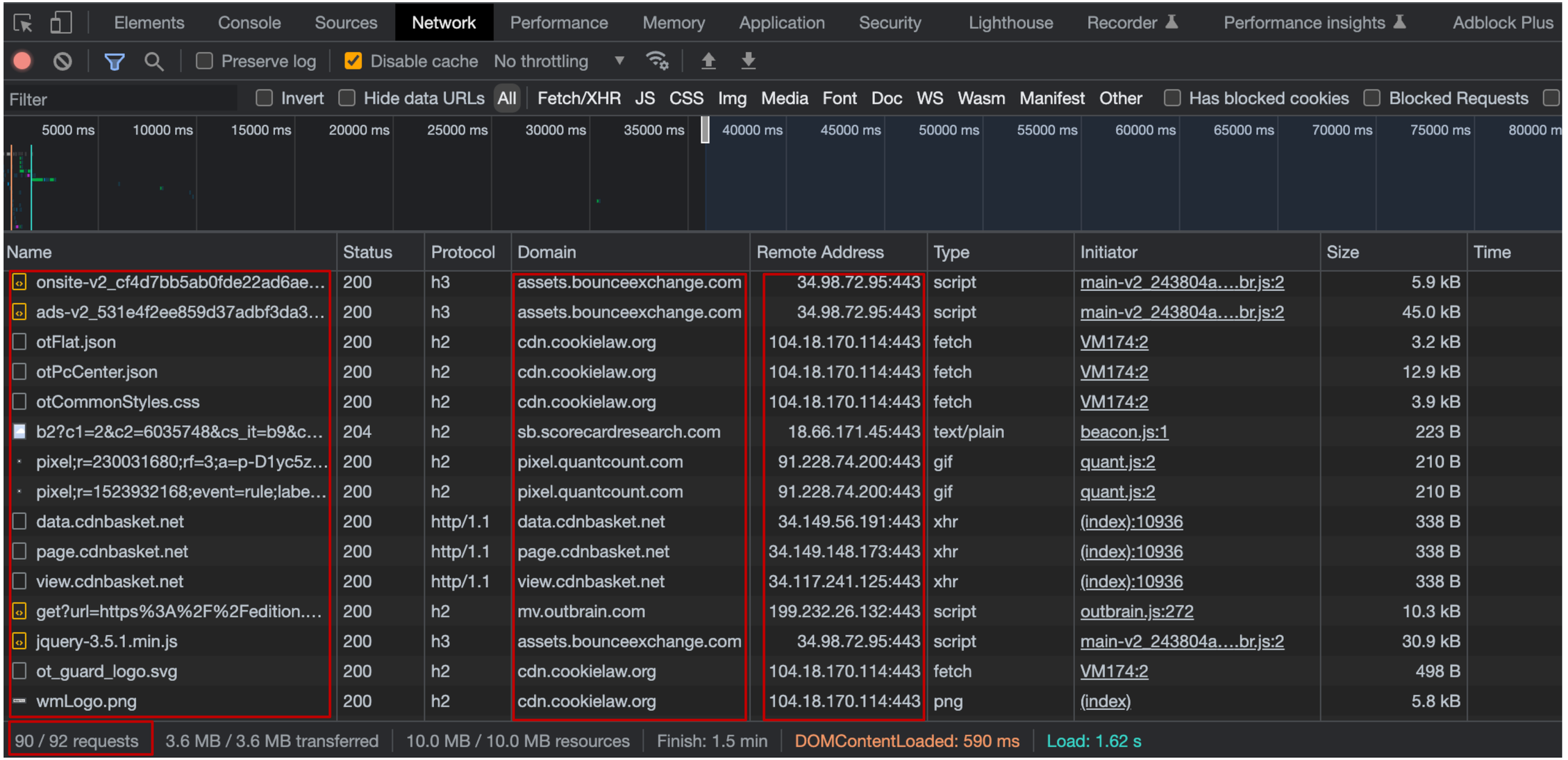


Anatomy of a Modern Web App



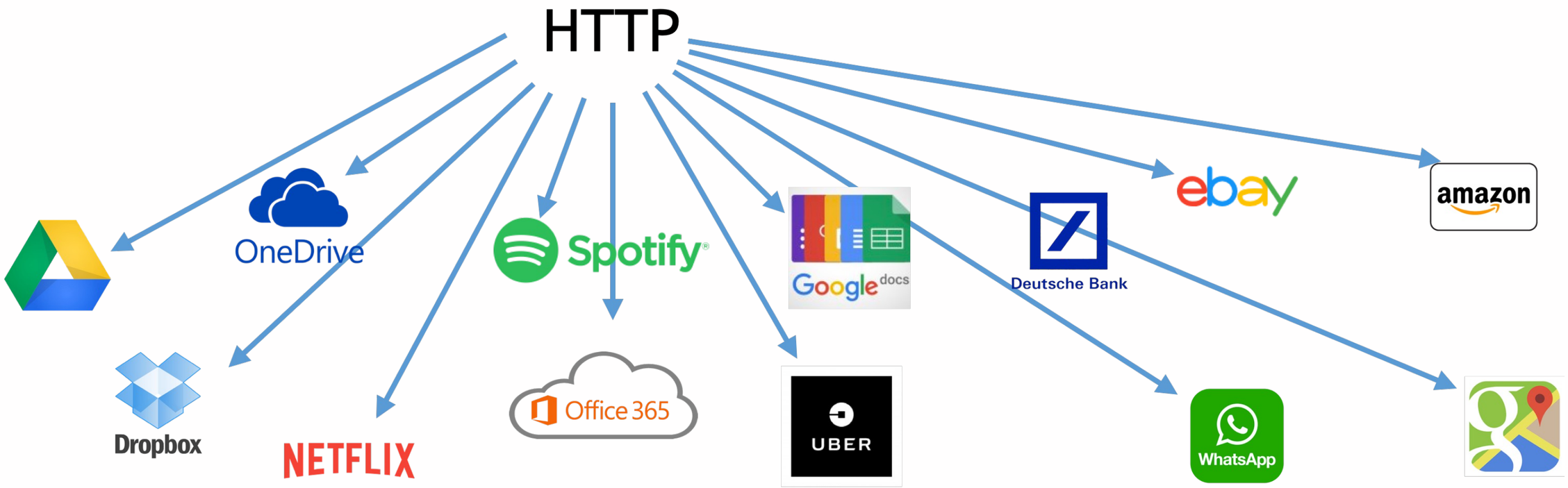
Anatomy of a Modern Web App

Ollscoil na Gaillimhe
University of Galway



How to Improve Web Performance?

- The complexity of web content over the years necessitated the need to update the Hypertext Transfer Protocol (HTTP) protocol.



HTTP History

Ollscoil na Gaillimhe
University of Galway

- HTTP is the foundation of data communication for the Web
- Originally designed for the transfer of simple web pages over the Internet in early 90’s
- https://home.cern/science/computing/birth-web/short-history-web
- info.cern.ch was the address of the world's first website and Web server, running on a computer at CERN.

Tim Berners-Lee, pictured at CERN
(Image: CERN)

HTTP History

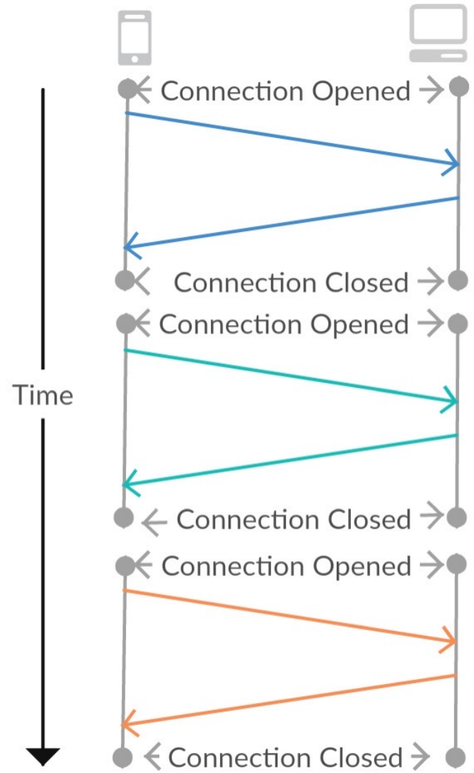
Ollscoil na Gaillimhe
University of Galway

- **HTTP/0.9**
  - In 1991, the first documented official version of HTTP was written as a plain document, less than 700 words long, and this version was named HTTP/0.9, which supported only GET method, allowing clients to only retrieve HTML documents from the server, but not supporting any other file formats or information upload.

HTTP 1.0

Ollscoil na Gaillimhe
University of Galway

- In May 1996, RFC 1945 was published as a final HTTP/1.0
- In HTTP 1.0, each request/response pair requires opening a new connection.
- **New Features:**
  - **Header:** The HTTP header was introduced, thus allowing the transmission of metadata that made the protocol flexible and extensible.
  - **Status code:** HTTP responses now contained a status code, thus enabling the receiver to check the request processing status (successful or failed).
  - **Content-type:** HTTP could transmit other documents types than a plain HTML file.
  - **New methods:** Two new methods POST and HEAD were provided, besides GET.

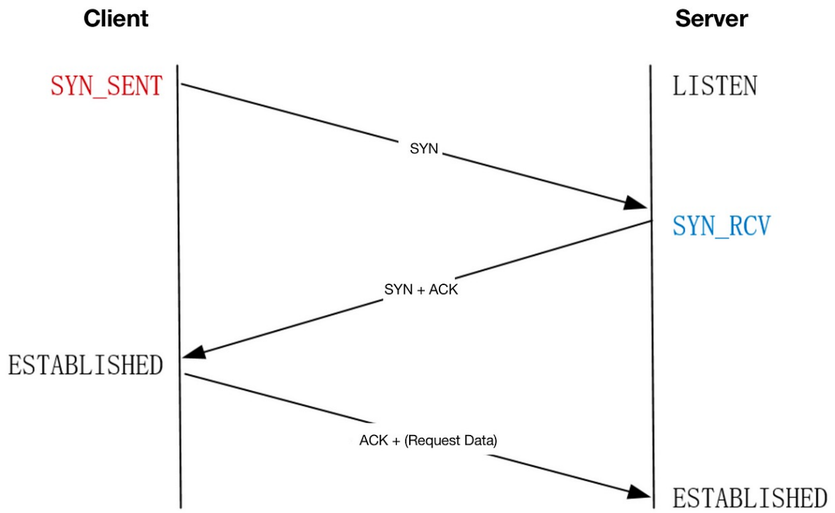


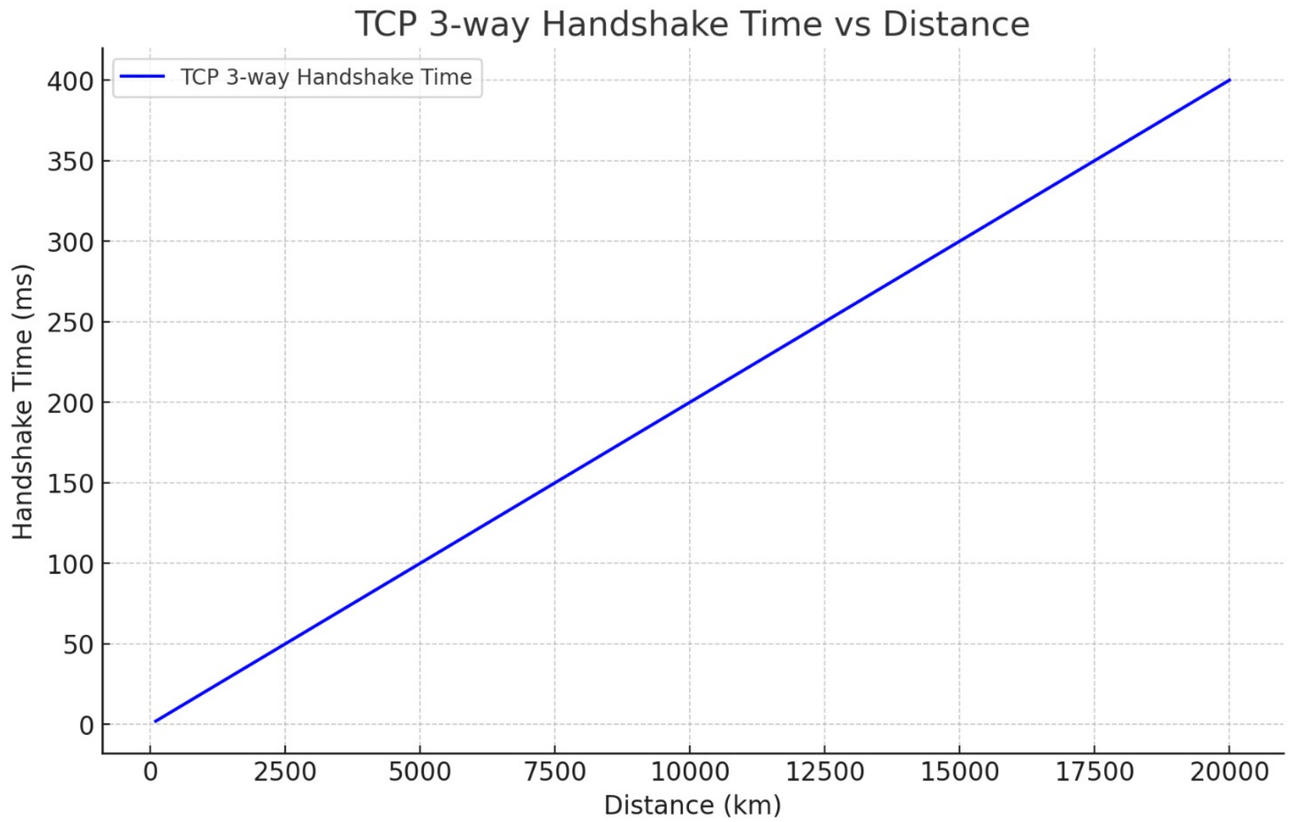
HTTP/1.0
Separate Connections

HTTP 1.0

Ollscoil na Gaillimhe
University of Galway

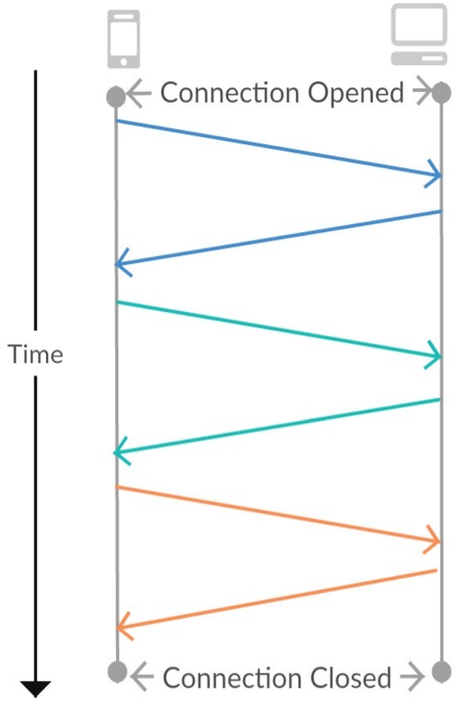
- Establishing a TCP connection using 3-way handshake is expensive.
- 2 RTTs between client and server are required to establish a connection.





HTTP 1.1

- In January 1997, RFC 2068 was officially released as HTTP/1.1 specifications.
- The following are the most relevant enhancements:
- **Persistent connections:** In HTTP 1.1, it is possible to execute several requests using a single connection, and thus amortize the cost of the initial connection establishment and slow start across multiple requests
- **New methods:** besides the already available methods of HTTP 1.0, the 1.1 version added six extra methods: PUT, PATCH, DELETE, CONNECT, TRACE, and OPTIONS



HTTP/1.1
Persistent Connections
Request Queuing

HTTP 1.1

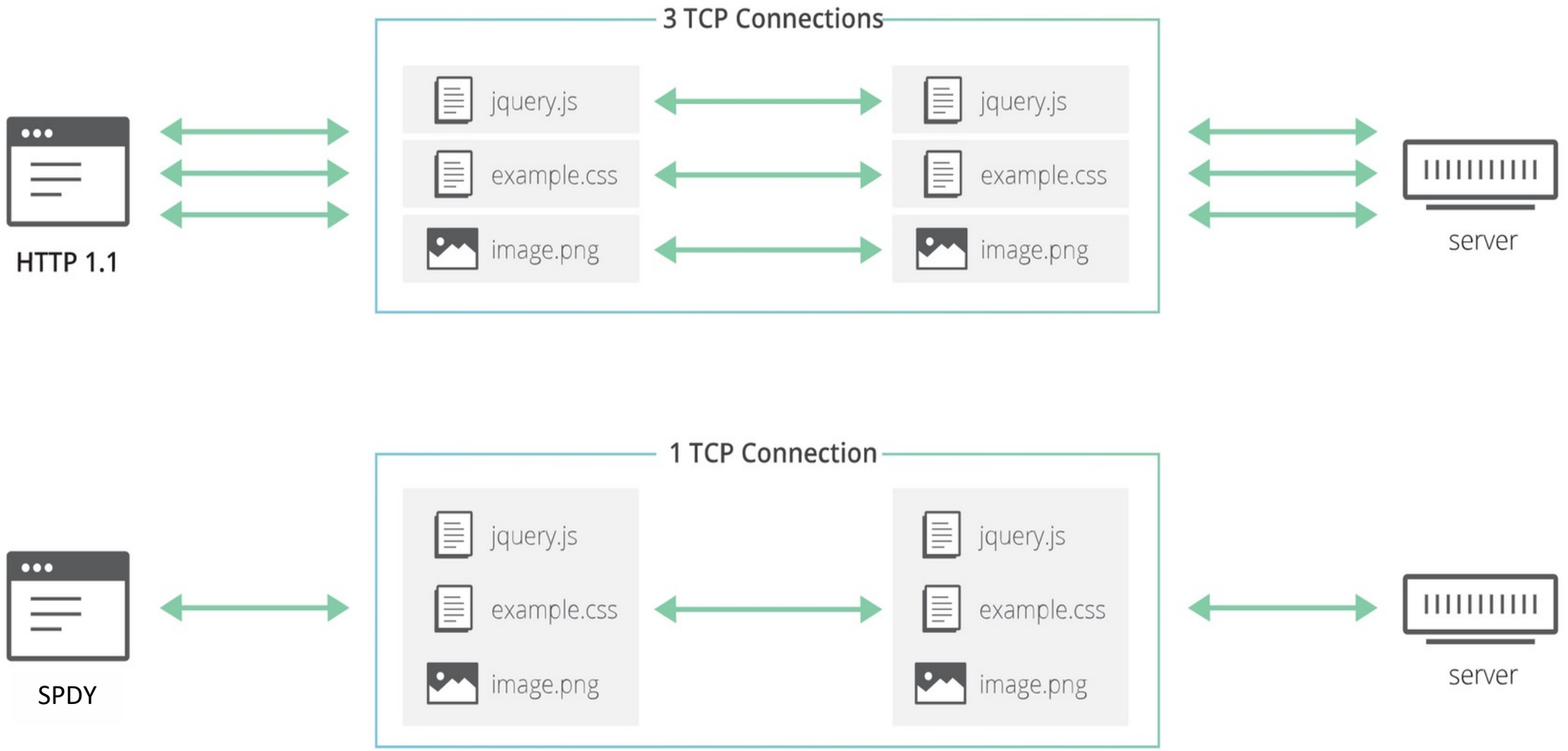
Ollscoil na Gaillimhe
University of Galway

- With persistent connections, multiple requests could share the same connection, but they still had to be serialized one after the other
  - So a client and server could only execute a single request/response exchange at any given time for each connection.
- As the web evolved, more concurrency was required when fetching and rendering web pages with a large number of resources (CSS, JavaScript, images, ...).
- The only way to gain concurrency at the network layer was to use multiple TCP connections to the same origin in parallel. But it has several negative effects.
  - Initial Overhead - Establishing and maintaining multiple connections requires additional overhead
  - Network Congestion - Multiple parallel connections increase the amount of traffic on the network
  - Inefficient Use of Resources - Every TCP connection requires resources on both the client and server sides, such as buffers, sockets, and memory
  - Fairness Issues - Some users or applications may monopolize bandwidth by opening many TCP connections

SPDY

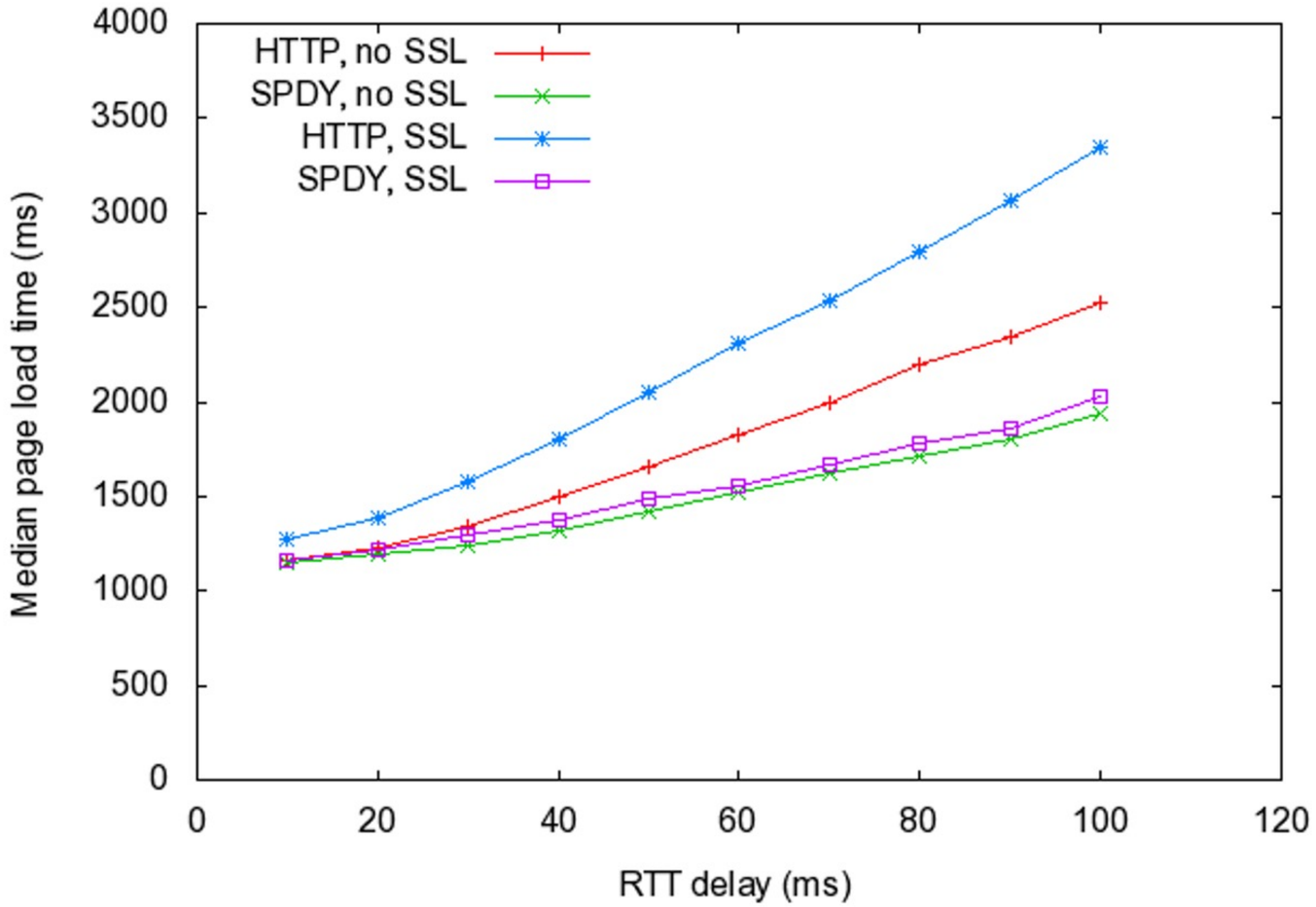
- An unofficial HTTP protocol developed by Google in 2009 as an experimental protocol to improve the web performance
  - SPDY opens one connection per domain
  - Multiple data streams are multiplexed over this single TCP connection for efficiency
  - Transfer higher priority resources faster than low priority resources
  - Reduces the amount of redundant header information each time a new page is requested.
- SPDY was supported by Chrome browser and deployed in most Google services

SPDY



https://blog.cloudflare.com/http3-the-past-present-and-future/

SPDY Performance



Page load time comparison for a dummy web page using 10Mbps connection

https://www.microsoft.com/en-us/research/publication/a-comparison-of-spdy-and-http-performance/ (2012)

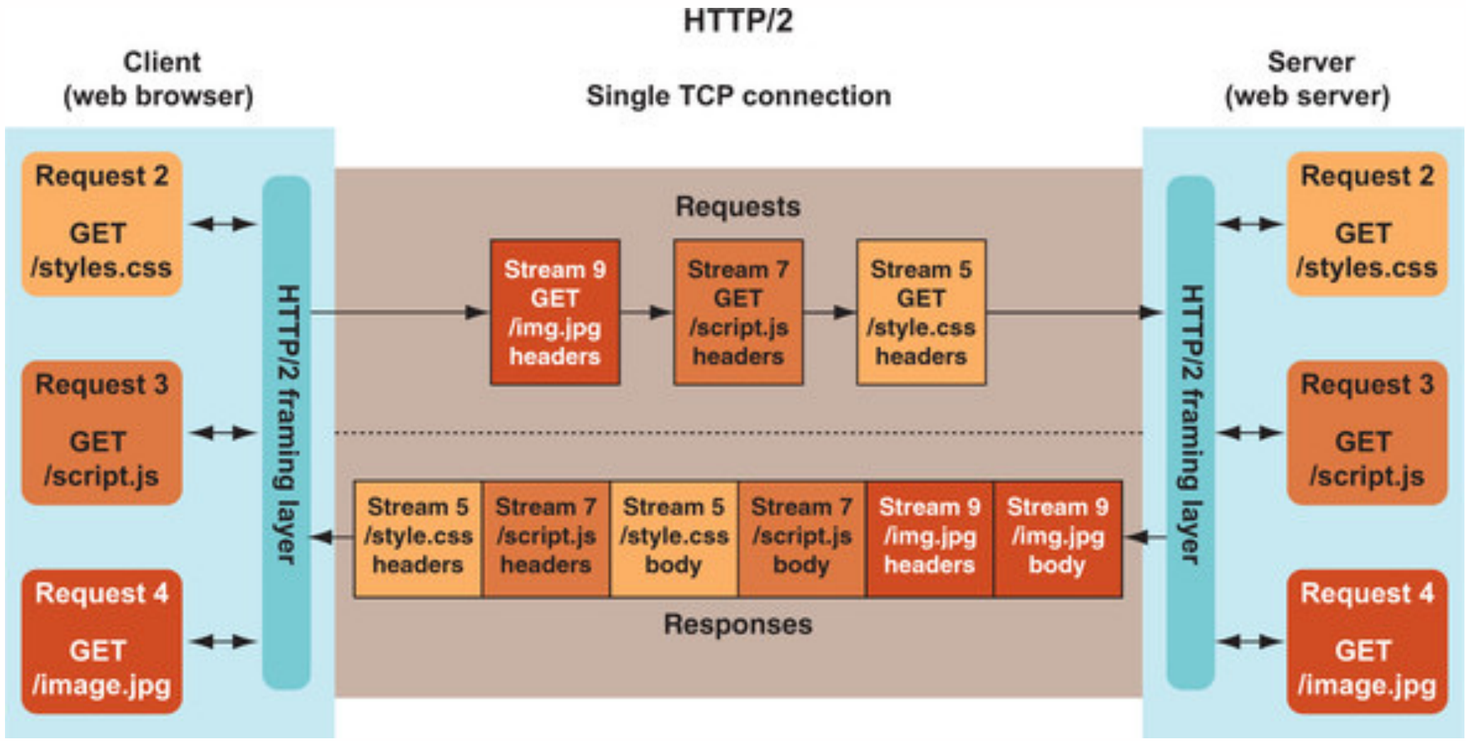
HTTP 2.0

Ollscoil na Gaillimhe
University of Galway

- It was officially released in 2015, (RFC 7540) about 18 years after the HTTP 1.1
- It was decided to derive it from SPDY
- HTTP 2.0 implemented several new features to improve the protocol performance:

HTTP 2.0 Features

1. **Request multiplexing:** HTTP 1.1 is a sequential protocol. So, we can send a single request at a time. HTTP 2.0, in turn, allows to send requests and receive responses asynchronously. In this way, we can do multiple requests at the same time using a single connection



http://manning.com/book/http2-in-action

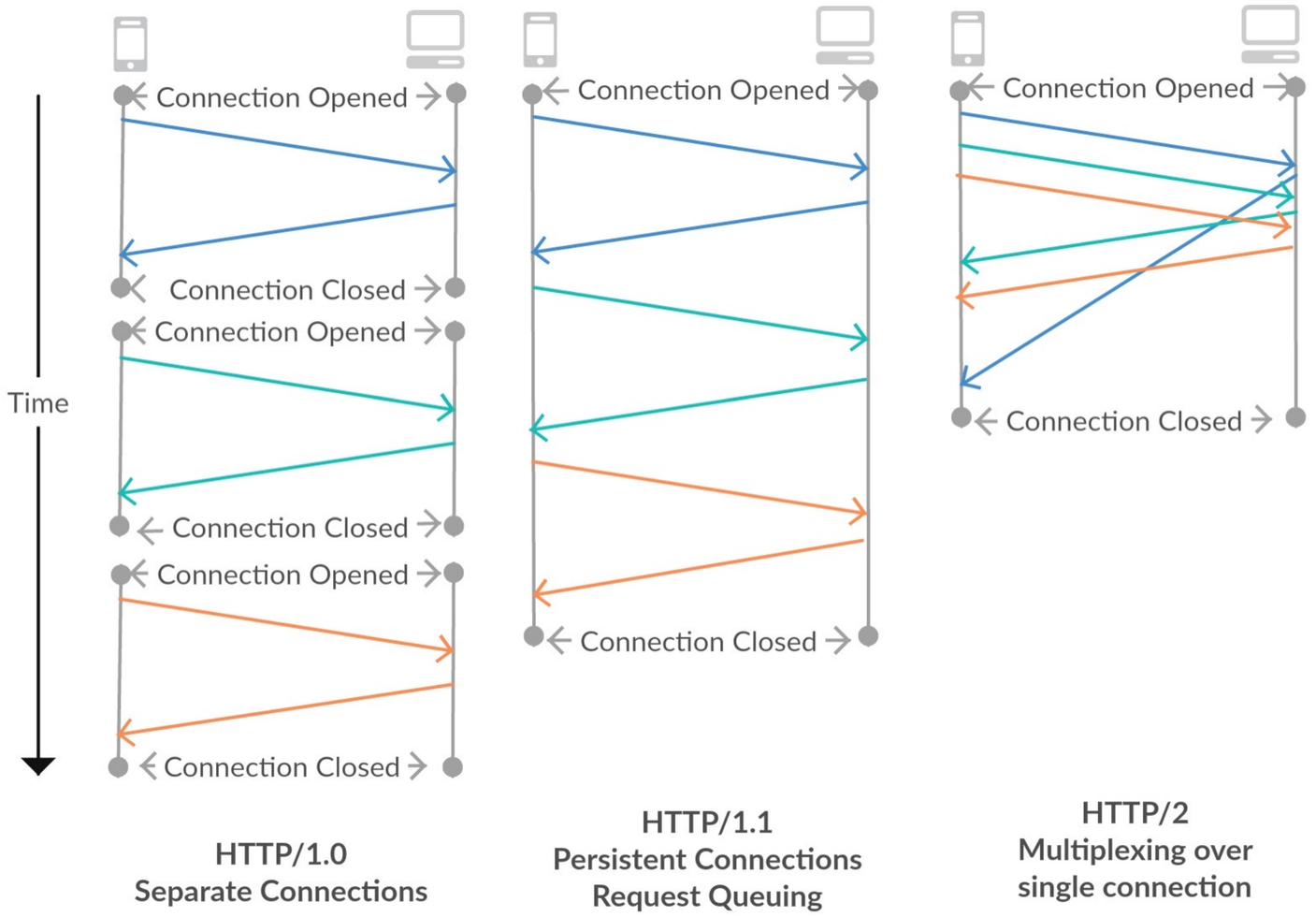
HTTP 2.0 Features

Ollscoil na Gaillimhe
University of Galway

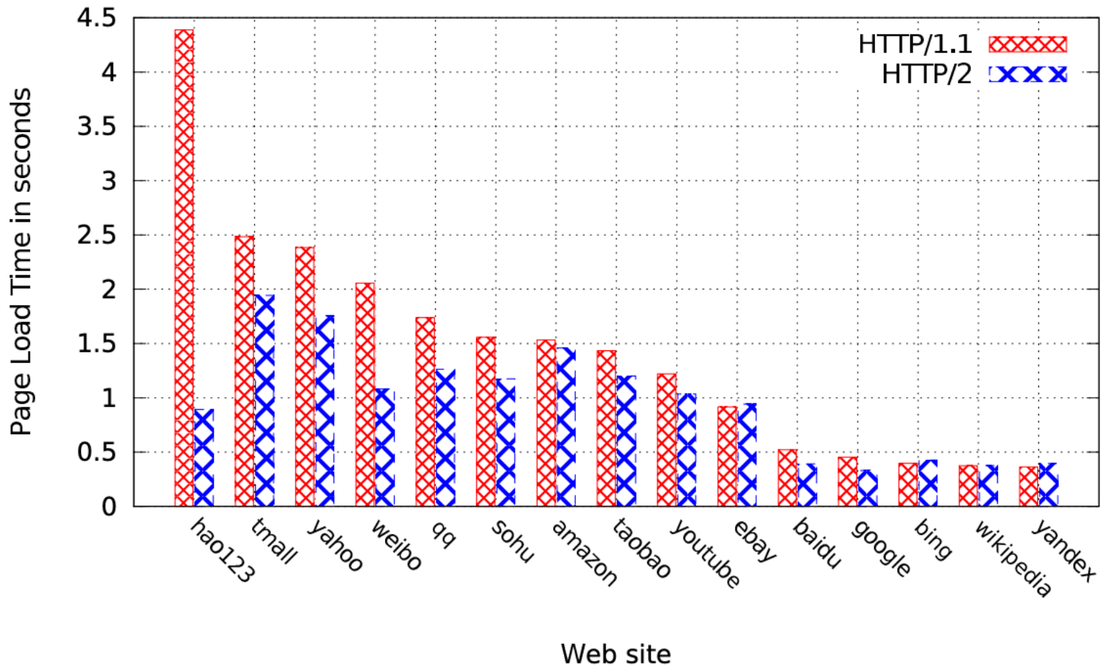
2. **Request prioritization:** We can set a priority for requests. Thus, we can be explicit in which order we expect the responses, such as getting a webpage CSS before its JS file
3. **Header compression:** Headers are compressed with HPACK (RFC 7541)
    - 76% compression of ingress header
    - 69% compression of egress header (https://blog.cloudflare.com/hpack-the-silent-killer-feature-of-http-2/)
4. **Server push:** to avoid a server receiving lots of requests, HTTP 2.0 introduced a server push functionality. With that, the server tries to predict the resources that will be requested soon. So, the server proactively pushes these resources to the client cache

HTTP version comparison

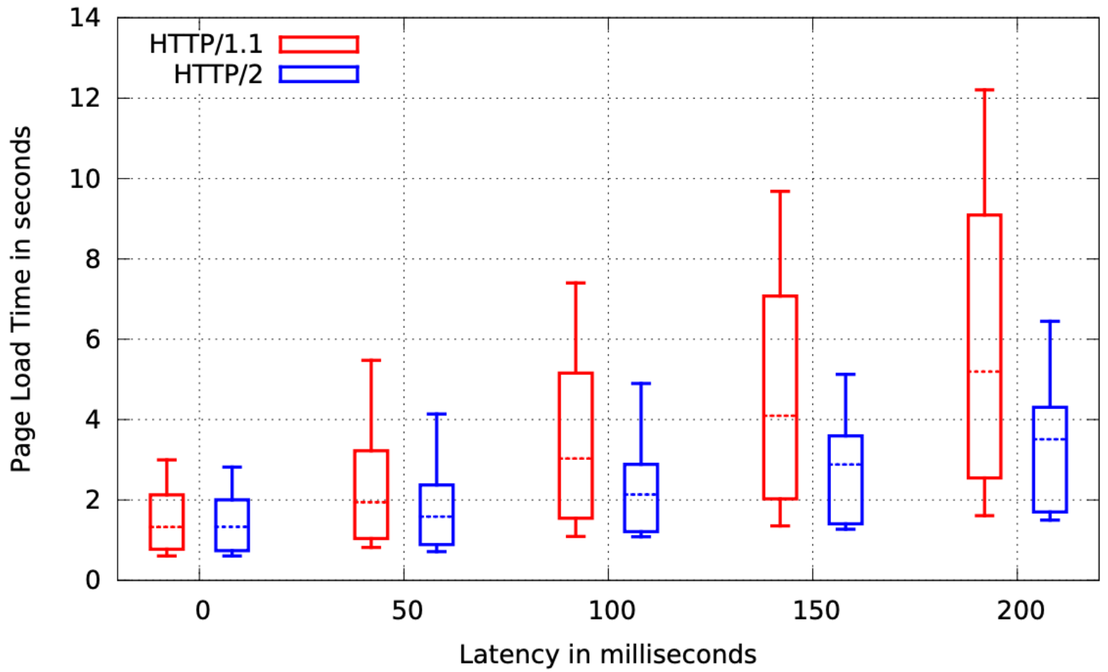
Ollscoil na Gaillimhe
University of Galway



HTTP 2.0 performance



Page load time for different websites using 50ms latency



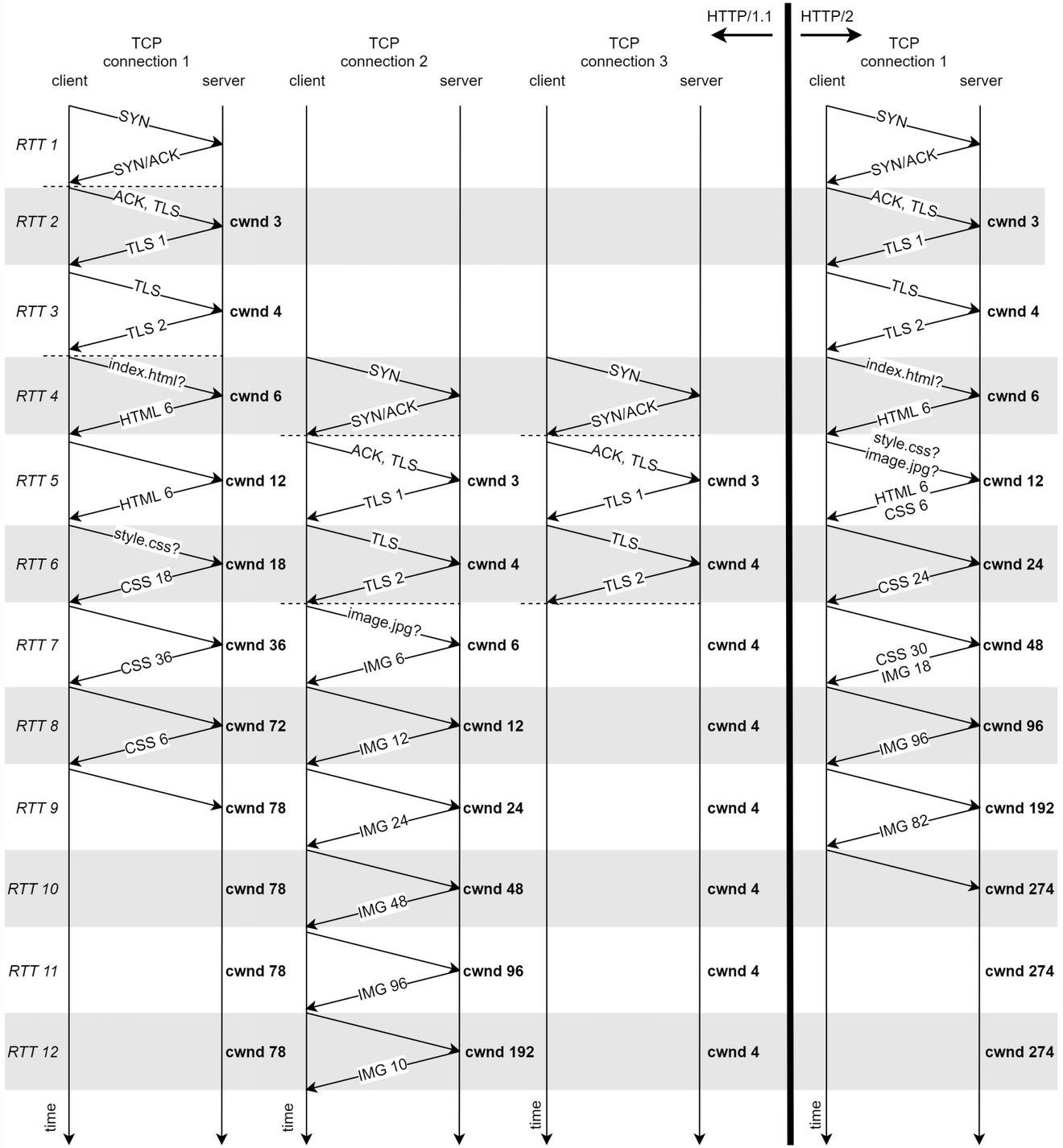
Impact of latency on page load time

"Is HTTP/2 really faster than HTTP/1.1?." *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2015.

Question

Ollscoil na Gaillimhe
University of Galway

- Compare HTTP2's multiplexing of multiple resources and HTTP/1.1 downloading multiple resources in parallel using multiple connections.
- Which of these two approaches has better performance?



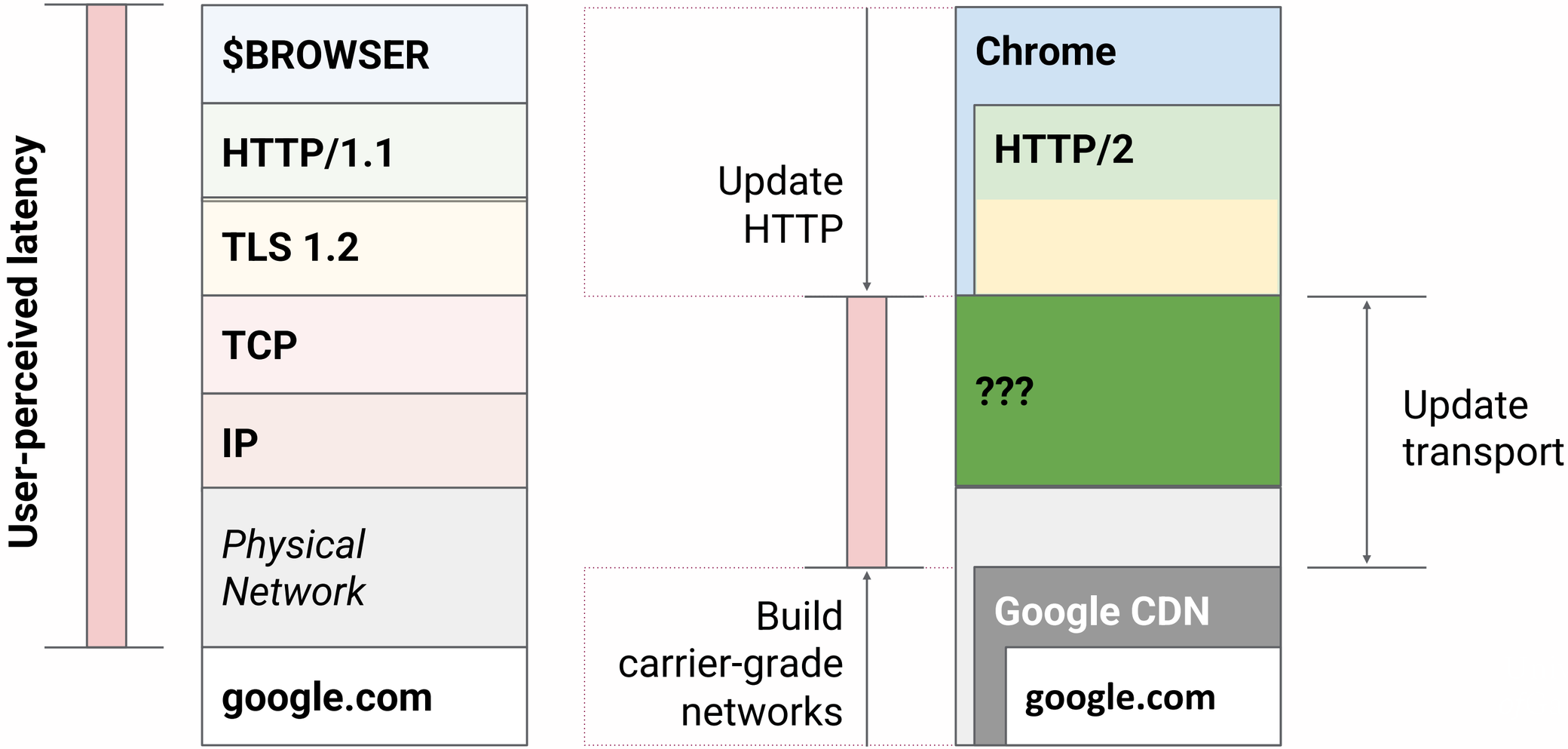
Ollscoil na Gaillimhe
University of Galway

Can we do better?

Ollscoil na Gaillimhe
University of Galway

- Improvements at the application layer have been implemented in HTTP 2.0
- To further improve the performance, fundamental changes to the underlying transport layer are required

Can we do better?



Ollscoil na Gaillimhe
University of Galway

Thank you for your attention!

University
*of*Galway.ie