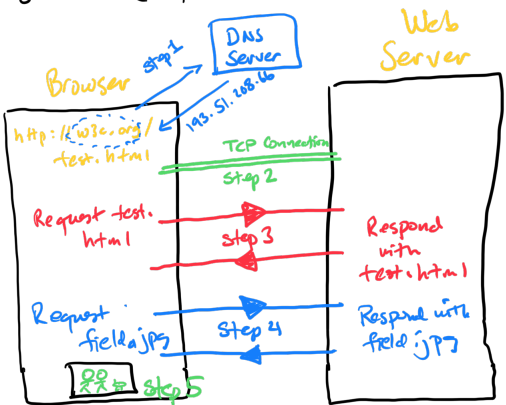# CPSC 216 - HyperText Transfer Protocol

→ HTTP works based on a request/response model.

→ The "client" application is typically a web browser (Chrome, Safari, ...)

→ The "server" application is a web server (Apache, ...)

Step 1: What does w3c.org mean?
- → browser extracts "domain name" from the URL
- → sends a request to the local domain name server, and that server returns the IP address of the web server.

Step 2: How do we talk to w3c.org?

- → browser establishes a TCP connection with the server, and prepares for information exchange

Step 3: Browser [requests] test.html
Web Server [responds] with test.html

```
<html>
<img src = "field.jpg">
</html>
```

Step 4: Browser "processes" the returned html file. It realizes that there is another required file that is needed.

Browser [requests] field.jpg
Web server [responds] with field.jpg

wee o...

Step 5: Browser displays image.

Step 6: TCP connection is closed?

OK, so let's dig into this a bit deeper...

Question: What actually gets sent in HTTP requests and responses?

Request for URL: https://en.wikipedia.org/wiki/World_Wide_Web
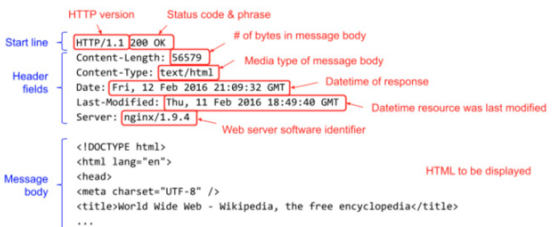
```
GET /wiki/World_Wide_Web HTTP/1.1
Host: en.wikipedia.org
User-Agent: Mozilla/5.0 Chrome/48.0.2564
```

Request type   Path   HTTP version

Start line { GET /wiki/World_Wide_Web HTTP/1.1

Header fields { Host: en.wikipedia.org ← Hostname
User-Agent: Mozilla/5.0 Chrome/48.0.2564

Browser identifier

Response for URL: https://en.wikipedia.org/wiki/World_Wide_Web

```
HTTP/1.1 200 OK
Content-Length: 56579
Content-Type: text/html
Date: Fri, 12 Feb 2016 21:09:32 GMT
Last-Modified: Thu, 11 Feb 2016 18:49:40 GMT
Server: nginx/1.9.4

<!DOCTYPE html>
<html lang="en">
<head>
<meta charset="UTF-8" />
<title>World Wide Web - Wikipedia, the free encyclopedia</title>
...
```

HTTP version   Status code & phrase

Start line { HTTP/1.1 200 OK
                                # of bytes in message body
Content-Length: 56579
Header   Content-Type: text/html ← Media type of message body
fields   Date: Fri, 12 Feb 2016 21:09:32 GMT ← Datetime of response
                Last-Modified: Thu, 11 Feb 2016 18:49:40 GMT ← Datetime resource was last modified
                Server: nginx/1.9.4
                                Web server software identifier

Message { <!DOCTYPE html>
body   <html lang="en">
           <head>   HTML to be displayed
           <meta charset="UTF-8" />
           <title>World Wide Web - Wikipedia, the free encyclopedia</title>
           ...

Each of the request and the response
have FOUR PARTS :

① Start Line
  Request: TYPE ⌴ path ⌴ HTTP version
  Response: HTTP version ⌴ CODE ⌴ phrase.
② Zero or more header fields
      keyword: value   (additional info)

③ Blank Line!

④ Optional Message Body

Common HTTP Request Methods:

GET   — request a resource
HEAD   — same as GET but no body
POST   — create new resource
PATCH   — modify a resource
PUT   — replace a resource
DELETE   — delete a resource

Common HTTP response status codes:

200 — OK   ✓
301 — Moved Permanently   —
302 — Found   —
304 — Not Modified   —
403 — Forbidden   —
404 — Not Found   ✗
500 — Internal Server Error   ✗