

# PART III

## RESEARCH PROJECT

BY AAKRITY PANDEY

2020MT60865

### HTTP

1.HTTP hypertext transfer protocol is a protocol for transmitting needed documents such as HTML documents . It is designed for communication between web browser and web servers. It works very typically .it sends request to the servers and waits until it receives a response . HTTP is a stateless protocol that means that server does not keep any data between 2 requests.

- i. HTTP sends requests from the client to server .The request comprises of 3 things ,start line, HTTP headers and body and there is an empty line between HTTP headers and the body. The start line comprises of 3 things HTTP methods request, target and the HTTP version. There are 3 elements of the headers request headers: general headers , representation headers the final part of the request is its body .
- ii. HTTP responses the responses sent by the server has a status line the headers and the body . The start line of the HTTP response is the status line which displays various codes like 404 ,203 ,402 indicating the status of failure or success of the request and other things. The body of HTTP responses comprises of the HTML documents and other contents that are to be communicated to the client.

2.The HTTP methods indicates the desired action that must be performed for the given resource that is to be accessed. There are various methods. The methods are only safe if they do not if they do not alter the state of the server.

Some HTTP methods are :

- ✓ GET: the get method is used to retrieve information from the given server using the given URL request. Using get should only retrieve data and have should no effect on that data.
- ✓ HEAD: Works same as get ,but transfers the status line and the header section only without the body .
- ✓ POST: post is used to send user information like username, the file the user is uploading or different things using HTML forms . E.g.: POST /cgi-bin/process.cgi HTTP/1.1/
- ✓ PUT: replace all the current representations of the target source with the uploaded content .This also alters the data on the server.
- ✓ DELETE: removes all current representations of the target URI source given by a URI.










3&4. USER AGENT: Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/90.0.4430.85 Safari/537.36 ( of Chrome browser on windows ).

Gecko is a layout engine or browser engine developed by Mozilla to convert HTML CSS JavaScript documents into what is presented to us on the browser screen.

History of user agent string

User agent : the user agent as the name suggests is a line of text displayed in the header that is used to identify the browser its version the details of the operating system etc. by the hosting web server.

The early web browser which popularised the Worldwide Web was Mosaic .

Browsers	User-Agent
 Mosaic was also the first browser to display images inline with text instead of displaying images in a separate window	<ul style="list-style-type: none"><li>- NCSA Mosaic/3.0 (Windows 95)</li><li>- NCSA_Mosaic/2.7b4 (X11;AIX 1 000180663000)</li><li>- NCSA_Mosaic/2.6 (X11; SunOS 4.1.3 sun4m)</li><li>- NCSA_Mosaic/2.0 (Windows 3.1)</li></ul>
	Mozilla/5.0 (Windows; U; Win 9x 4.90; SG; rv:1.9.2.4) Gecko/20101104 Netscape/9.1.0285
	Mozilla/5.0 (Windows NT 6.1; WOW64; Trident/7.0; AS; rv:11.0) like Gecko
	Mozilla/5.0 (Windows; U; Windows NT 6.1; rv:2.2) Gecko/20110201
	Mozilla/5.0 (Windows NT 5.2; RW; rv:7.0a1) Gecko/20091211 SeaMonkey/9.23a1pre
	Mozilla/5.0 (X11; Linux) KHTML/4.9.1 (like Gecko) Konqueror/4.9
	Opera/9.80 (X11; Linux i686; Ubuntu/14.10) Presto/2.12.388 Version/12.16
	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_9_3) AppleWebKit/537.75.14 (KHTML, like Gecko) Version/7.0.3 Safari/7046A194A
	Mozilla/5.0 (Windows NT 6.1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/41.0.2228.0 Safari/537.36

Then came Mozilla which meant Mosaic killer. The Mosaic was not amused by this and therefore Mozilla change its public name to Netscape. The Mosaic only displayed photos along with text .They did not support frames. Therefore, the user agent sniffing started to send out frames only to Netscape browser . Then came Internet Explorer by Microsoft which supported frames but was unable to get the pages designed for Netscape browser. That is why they decided to start its user agent string with Mozilla. Therefore, the websites which opened in Mozilla could be open in its browser too. This shortcut started the chaos in user agent string, the way to identify Mozilla and Internet Explorer was that in the middle of the user agent string somewhere they inserted the version number which differentiated the 2 browsers .

Explorer Internet Explorer took over Netscape for a while, but Netscape launched its new browser Mozilla with a new functionality and rendering engine called Gecko built in it. Then Mozilla was renamed as Firefox and all of them the new browsers invented pretended to be Mozilla and were powered by gecko. Then on as they window operating systems and web browser , operating system got updated the user agent string became more and more complicated.

# COOKIES

1.Cookies are text files with small piece of data like username and password that are used to identify you as you ,on the computer network. When cookie is exchanged between you and your computer network the server reads your ID and knows what information to display to you without asking.

Cookies are essential to modern Internet but a vulnerability to your privacy. As a part of web developing the HTTP cookies enable web developers to remember your personal information like your website logins, your visited pages, your shopping carts and the ads which you liked and so on this way they are able to give you a more personalised experience in the website browsing.

2.The different attributes that can be applied while setting a cookie are :

✓ Secure attribute

The secure attribute tells the browser to only send cookie if the request is being sent over a secure channel such as HTTPS. This will protect the cookie from being sent to unencrypted requests.

✓ HTTP only attribute

It prevents attacks such as session leakage as it doesn't allow cookie to be accessed by a client side script such as JavaScript.

✓ Domain attribute

The domain attribute is used to compare the cookies domain against the domain of the server for which HTTP request is being made if the domain matches or if it is a sub domain the path attribute will be checked .Next now that only hosts that belong to specified domain can set a cookie for that domain. Additionally, the domain attribute cannot be a top-level domain (such as .gov or .com ) to prevent servers from setting arbitrary cookies for another domain. If the domain attribute is not set, then the host name of the server generated the cookie is used as a default value in the domain.

✓ Path attribute

The path attribute sets of the scope of the cookie in conjunction with the domain.

3.Advertising companies such as Google uses 3rd party cookies that are placed on other websites ,shopping websites like flipkart, amazon we visit and the data is sent to Google so that they can display ads accordingly .

Moreover,

A Flash cookie is a small file stored on your computer by a website that uses Adobe's Flash player technology. Flash cookies use Adobe's Flash player to store information about your online browsing activities.

4.

The main types of cookies that create privacy concerns for users are third party cookies. Although first-party cookies are cookies sent by the website you have entered, third-party cookies are cookies that are created and placed on your device by different internet subjects placed on the website you are visiting.

Federated Learning of Cohorts (FLoC) proposes a new way for businesses to reach people with relevant content and ads by clustering large groups of people with similar interests. This approach effectively hides individuals “in the crowd” and uses on-device processing to keep a person’s web history private on the browser.

By creating simulations based on the principles defined in Chrome’s FLoC proposal, Google’s ads teams have tested this privacy-first alternative to third-party cookies. Results indicate that when it comes to generating interest-based audiences, FLoC can provide an effective replacement signal for third-party cookies. Our tests of FLoC to reach in-market and affinity Google Audiences show that advertisers can expect to see at least 95% of the conversions per dollar spent when compared to cookie-based advertising. The specific result depends on the strength of the clustering algorithm that FLoC uses and the type of audience being reached.

Others worry that FLoC is just Google attempting to dress up what ostensibly is at its core another, albeit potentially less obtrusive way to track people’s behavior to suit its targeted advertising agenda to ensure the company will continue to drive the market.

“Google has announced that its tests show promising signs that FLoC is working,” wrote Malwarebytes Labs security research Pieter Artnz in a blog post published in January. “Is this a milestone on the road to more privacy, or just better concealed tracking technology?”

# CORS

Cross-Origin resource sharing is a header based mechanism that allows the server to indicate any other origin than its own for which a browser should permit loaded of resources .

Other origin means the URL being accessed differs from the location that the JavaScript is running from by having:

A different scheme

A different domain

A different port

Ports mechanism relies on sending pre-flight requests to this server. The HTTP request OPTIONS method is sent to the resource on the other origin in order to determine if the actual request is safe to send across site requested site pre-flighted like this since they may have implications to data.

```
OPTIONS /doc HTTP/1.1
```

```
Host: bar.other
```

```
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10.14; rv:71.0) Gecko/20100101  
Firefox/71.0
```

```
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
```

```
Accept-Language: en-us,en;q=0.5
```

```
Accept-Encoding: gzip,deflate
```

```
Connection: keep-alive
```

```
Origin: http://foo.example
```

```
Access-Control-Request-Method: POST
```

```
Access-Control-Request-Headers: X-PINGOTHER, Content-Type
```

```
HTTP/1.1 204 No Content
```

```
Date: Mon, 01 Dec 2008 01:15:39 GMT
```

```
Server: Apache/2
```

```
Access-Control-Allow-Origin: https://foo.example
```

```
Access-Control-Allow-Methods: POST, GET, OPTIONS
```

```
Access-Control-Allow-Headers: X-PINGOTHER, Content-Type
```

```
Access-Control-Max-Age: 86400
```

```
Vary: Accept-Encoding, Origin
```

```
Keep-Alive: timeout=2, max=100
```

```
Connection: Keep-Alive
```

The first part represents the pre-flight request response.

```
Access-Control-Request-Method: POST
```

```
Access-Control-Request-Headers: X-PINGOTHER, Content-Type
```

Access the access control request method header notifies the server as a part of pre-flight request that when the actual request is sent it will be a post method in this way the server determines whether it wishes to accept the request under the circumstances or not. Only once the pre flight request is accepted is the actual request sent .