Web Architecture

Farooq Ahmed, FAST-NU, Lahore

Questions

- What is world-wide web?

- What is Internet?

- Are www and Internet same or different?

- How the web works?

- What it takes to program web?

- What issues warrant a web developers attention?

Internet

- **Internet** is a global network of computers communicating through **internet** over diverse underlying infrastructure

- The core enabler of internet is TCP/IP protocol suite worked out by Robert Kahn and Vint Cerf in 1973

- Early contributing organizations:
    - DARPA,
    - Stanford Research Institute
    - CERN

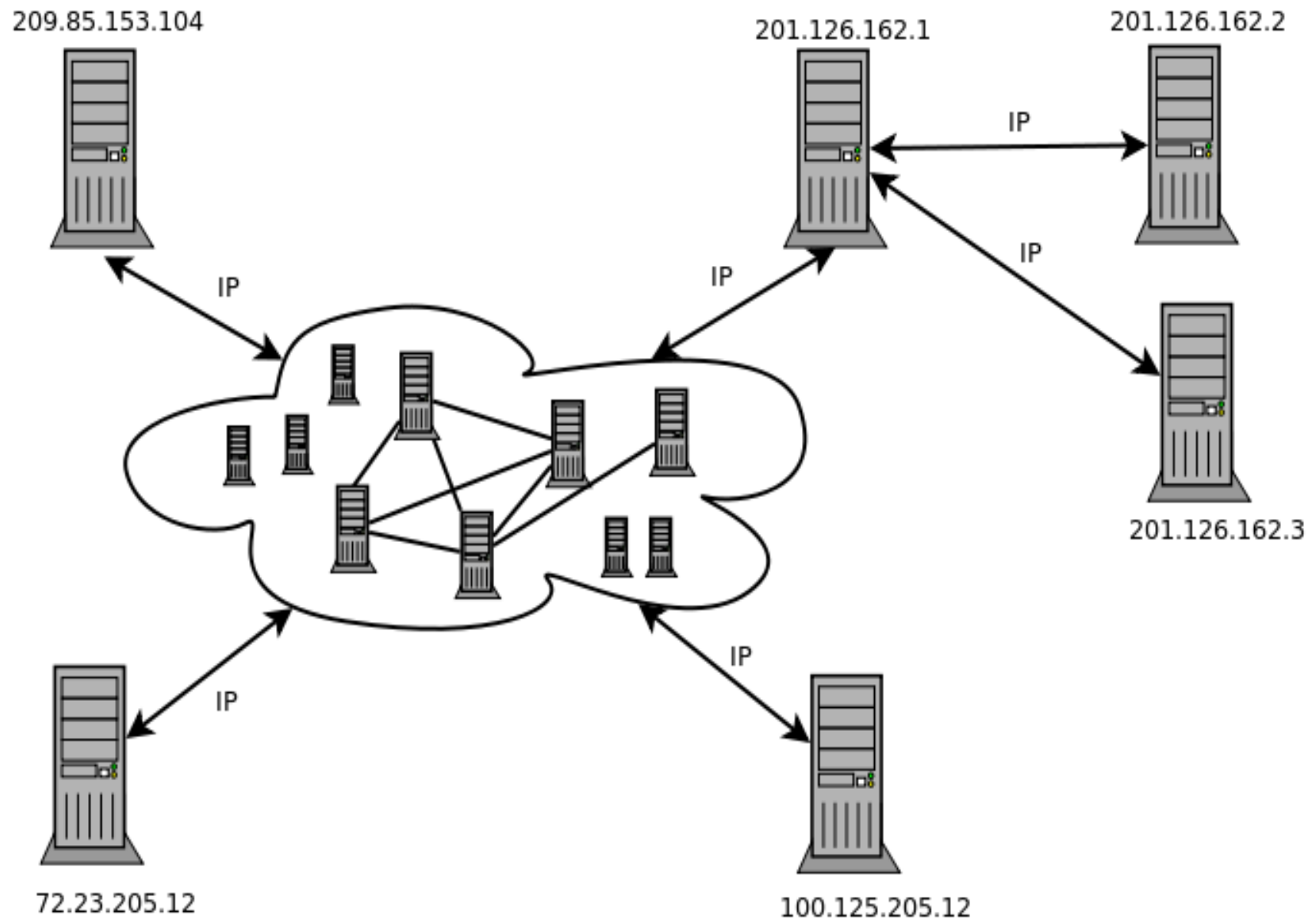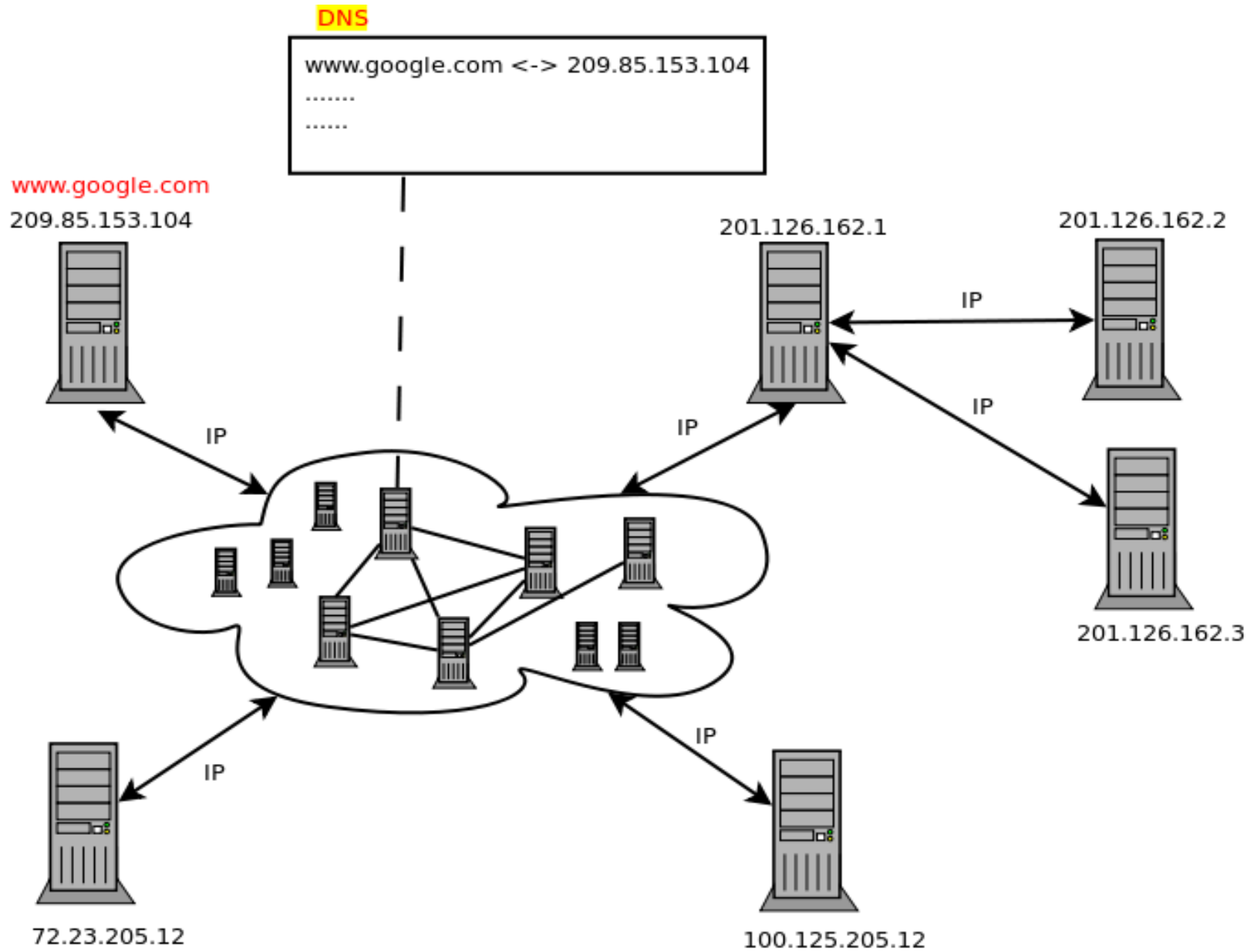- Current organizations:
    - IETF
    - ICANN

Robert Kahn
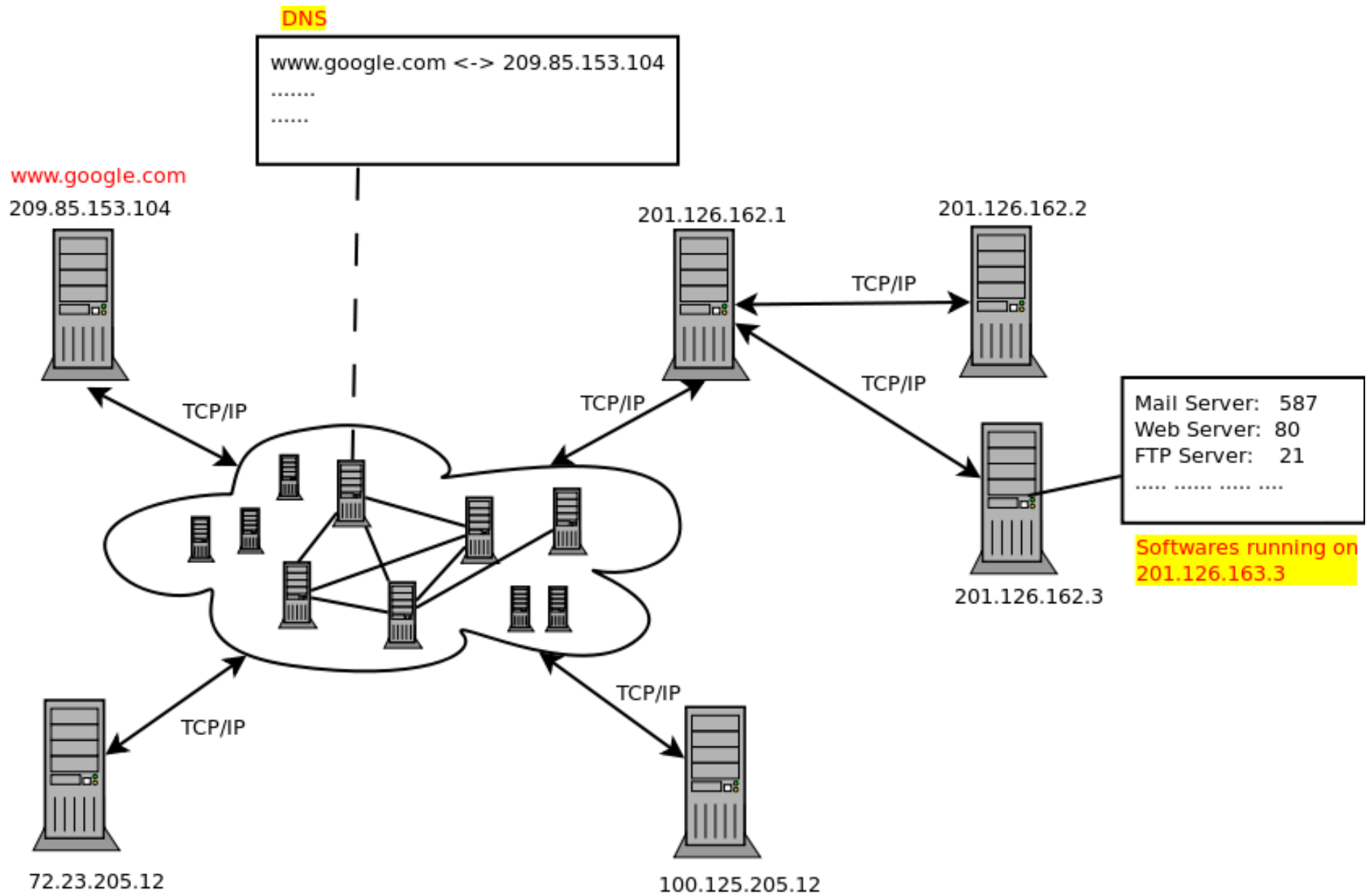Picture source: Wikipedia

Vinton Cerf

Farooq Ahmed, FAST-NU, Lahore

209.85.153.104

201.126.162.1

201.126.162.2

IP

IP

IP

IP

201.126.162.3

IP

IP

72.23.205.12

100.125.205.12

Farooq Ahmed, FAST-NU, Lahore

DNS

www.google.com <-> 209.85.153.104
....
......

www.google.com
209.85.153.104

201.126.162.1

201.126.162.2

IP

IP

IP

IP

IP

201.126.162.3

IP

IP

IP

72.23.205.12

100.125.205.12

Farooq Ahmed, FAST-NU, Lahore

www.google.com <-> 209.85.153.104
....
......

www.google.com
209.85.153.104

201.126.162.1

201.126.162.2

TCP/IP

TCP/IP

TCP/IP

TCP/IP

Mail Server: 587
Web Server: 80
FTP Server: 21

....

201.126.162.3

TCP/IP

TCP/IP

72.23.205.12

100.125.205.12

Farooq Ahmed, FAST-NU, Lahore

internet-based services

- File transfer and sharing (FTP)
- Remote login and management (Telnet, SSH, etc)
- Email (SMTP)
- Voice over IP (Skype)
- Virtual Private Networks (VPN)
- **World-wide web (HTTP)**
- …

World-wide Web (WWW)

- WWW is a service used world over for information sharing

- Introduced in 1990 by Tim Berners Lee

- One of the services available over Internet

- Comprises of a client-server architecture

- Three major components
  - Web Server
  - Agent / Browser
  - HTTP Protocol

- Early contributing organization
  - CERN

- Current contributing organization
  - W3C

Tim Berners Lee
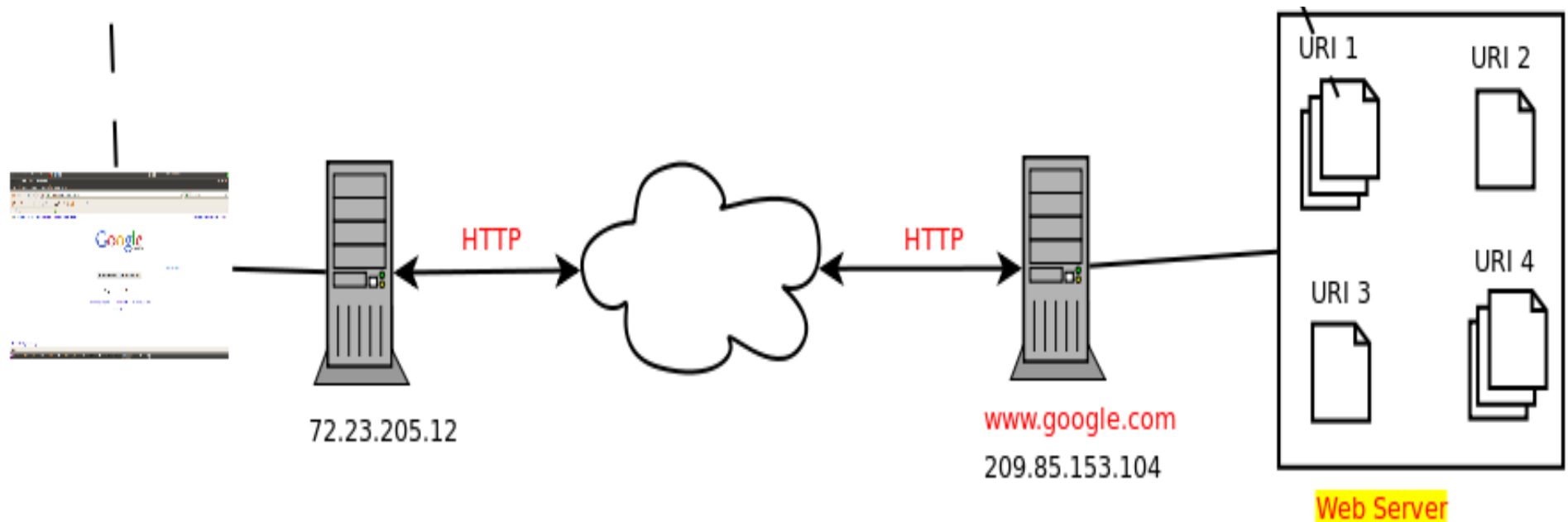Picture source: Wikipedia

WWW as service over Internet

Farooq Ahmed, FAST-NU, Lahore
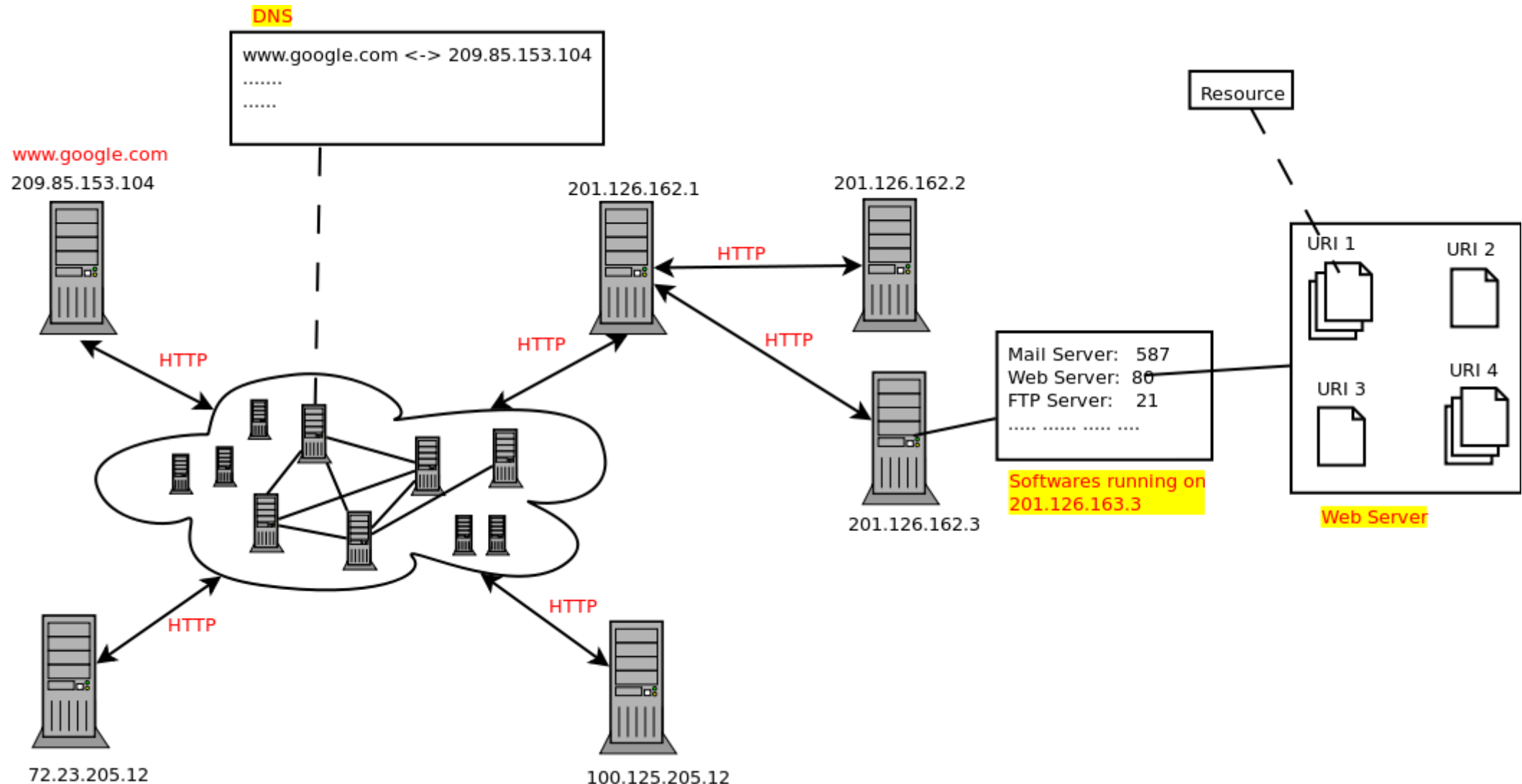
WWW as client-server architecture

Browser as a (thin) client

- Web server holds web content
- Browser presents content to end user
- HTTP is used to pull content from server to browser

Farooq Ahmed, FAST-NU, Lahore

Web Server

- ## Web server job is to:

  - host web content (static / dynamic)
  - Support technologies for efficient creation, deployment and distribution of content
  - Support multiple concurrent user requests
  - Support secure communication
  - Support multiple websites / applications simultaneously

- ## Modern Web servers

  - Apache
  - IIS
  - …

Browser
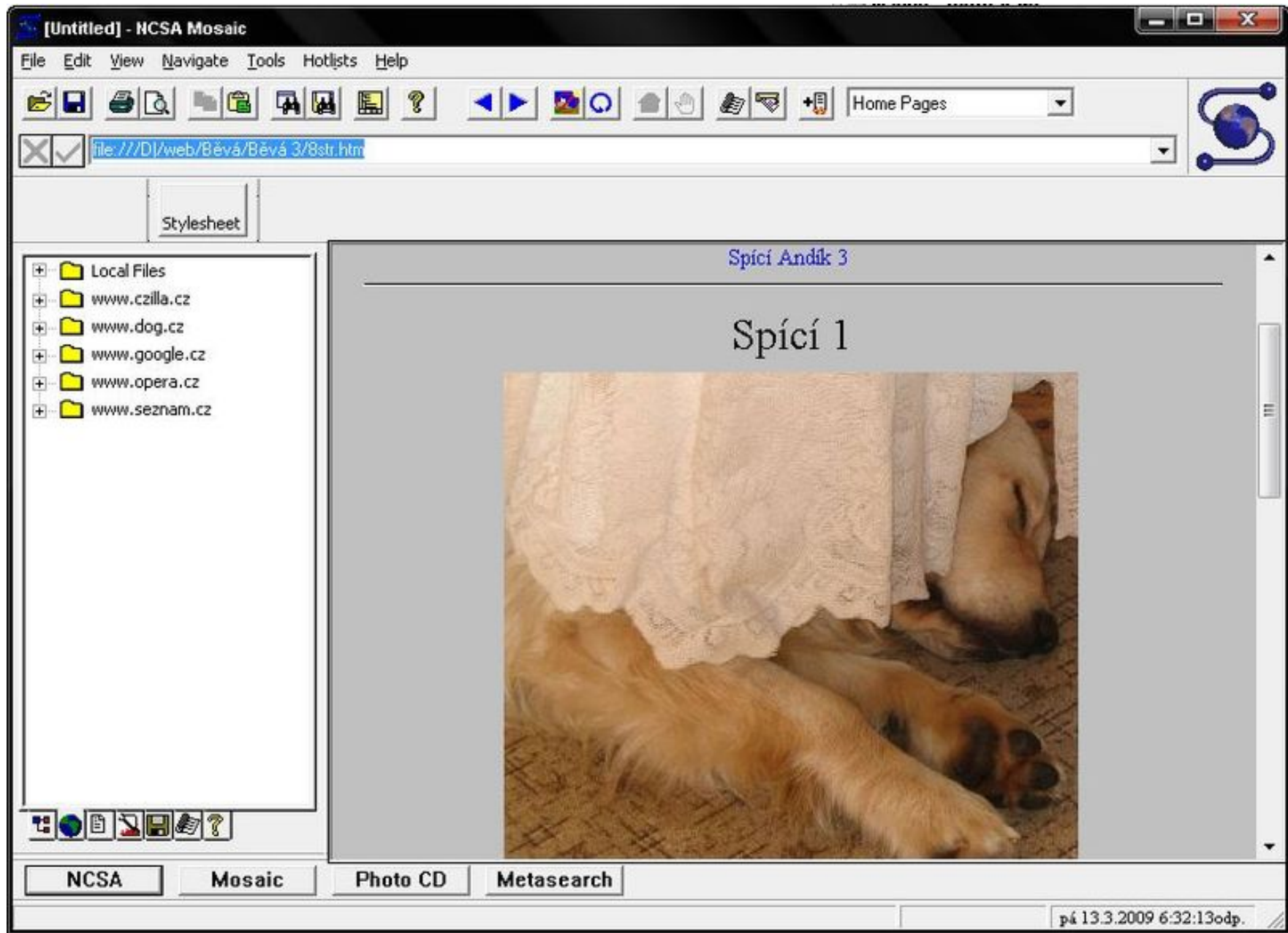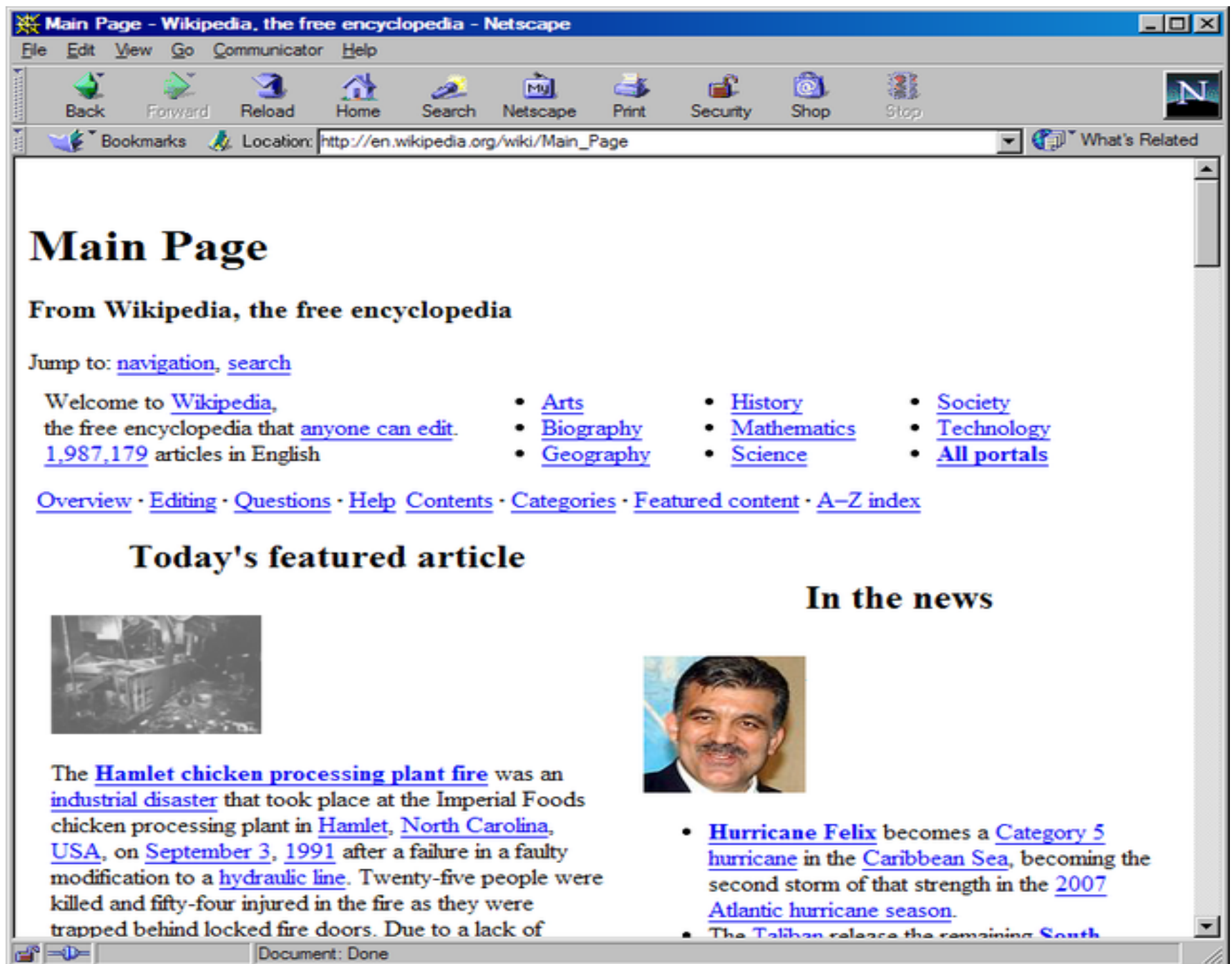

Mark Andreesen
Picture source : Wikipedia

- Browser function is to:
  - Fetch web content
  - Present web content
  - Efficient rendering
  - Secure communication
  - Support extensibility
- Early browsers
  - WorldWideWeb
  - Mosaic
  - Netscape Navigator
  - …
- Modern browsers
  - Mozilla Firefox
  - Google Chrome
  - Microsoft Internet Explorer
  - …

WorldWideWeb
Info ▷
Navigate ▷
Document ▷
Edit ▷
Links ▷
Style ▷
Print... p
Page layout...
Windows ▷
Services ▷
Hide h
Quit q

Style ✕
Copy style 1
Apply style 2
Panel...

Info ✕

HyperMedia Browser/Editor

Prerelease b of Version 0.13

An excercise in global information availability

by Tim Berners-Lee

Copyright 1990,91, CERN. Distribution restricted: ask for terms. TEST VERSION ONLY

HyperText: Text which is not constrained to be linear.
HyperMedia: Information which is not constrained linear... or to be text.

This version of the WWW application can pick up hypertext infomation from files in a number of formats, from local files, from remote files using NFS or anonymous FTP, from hypertext servers by name or keyword search, and from internet news. Hypertext files may be edited, and links made from hypertext files to other files or any other information.

For more help, use "Help" from the menu. If that doesn't work, then your application has been incompletely installed.

If you have any comments or have bugs, please mail timbl@info.cern.ch. quoting the version number (above).

Navigation ✕
< Previous   Back up   Next >
Home   Help

JFG Home ✕

JFG's home page at InfoDesign

This home page demonstrates some simple concepts of the World-Wide Web infrastructure for global information sharing.

Today, I showed Tim's original Web browser to Marc Weber. I just took some notes about him and linked them to his name on the fly.

I use this World-Wide Web editor exactly like any word processor, with the power to create links from sensitive pieces of text to other places, for instance to personal notes or to any information source in the world.

WWW research stuff

- Hypertext Resources (IN-Box)
- Geographical list of WWW servers (at C same list. Example : CERN phone book
- WWW project documentation (at CERN

Links ✕
Mark all   A
Mark selection   M
Link to marked   L
Link to file...
Link to New   N
Follow link
Unlink   Z
Help

Style editor

Style name: List

<<   style of selection   >>

Apply style to selection

Apply style to all similar text

Find unstyled text

Style sheet
Open
Save as...

Format
First line indent: 100
Successive indent: 130
Font: Helvetica
Point Size: 12
SGML tag: <LI>   Set
Tabs: 130 300

HotWired: What's New

Welcome to Yorb ,   the electr neighborhood hits Wired, 24 May. via in On the R

Flux Ned hears rumors of a deal between GN

Fetish Solar

Net Surf Pro

Jenny Holzer's truisms take a turn for the We

If you feel a need for speed ... surf t

What's a week without a Bill Gates j

Corporate recruiting form letters = ti

Alt.cynicism ... What's the Point?

free.

Weber -- /Hypertext

Marc Weber

A computer consultant, small publisher and freelance journalist who is currently investigating the early history of the Web, to be published in Wired. He loves HotWired.

Welcome To HotWired! ✕

How to Join: Register Now   Members Only:
Unguided Tour   Overview   Need Help?   What's New
Your View © 1995 HotWired Ventures LLC

Powered by Silicon Graphics

NeXT

WorldWideWeb browser by Tim Berners Lee developed at CERN (source : Wikipedia)

Mosaic browser by Mark Andreesen developed at UIUC (source : Wikipedia)
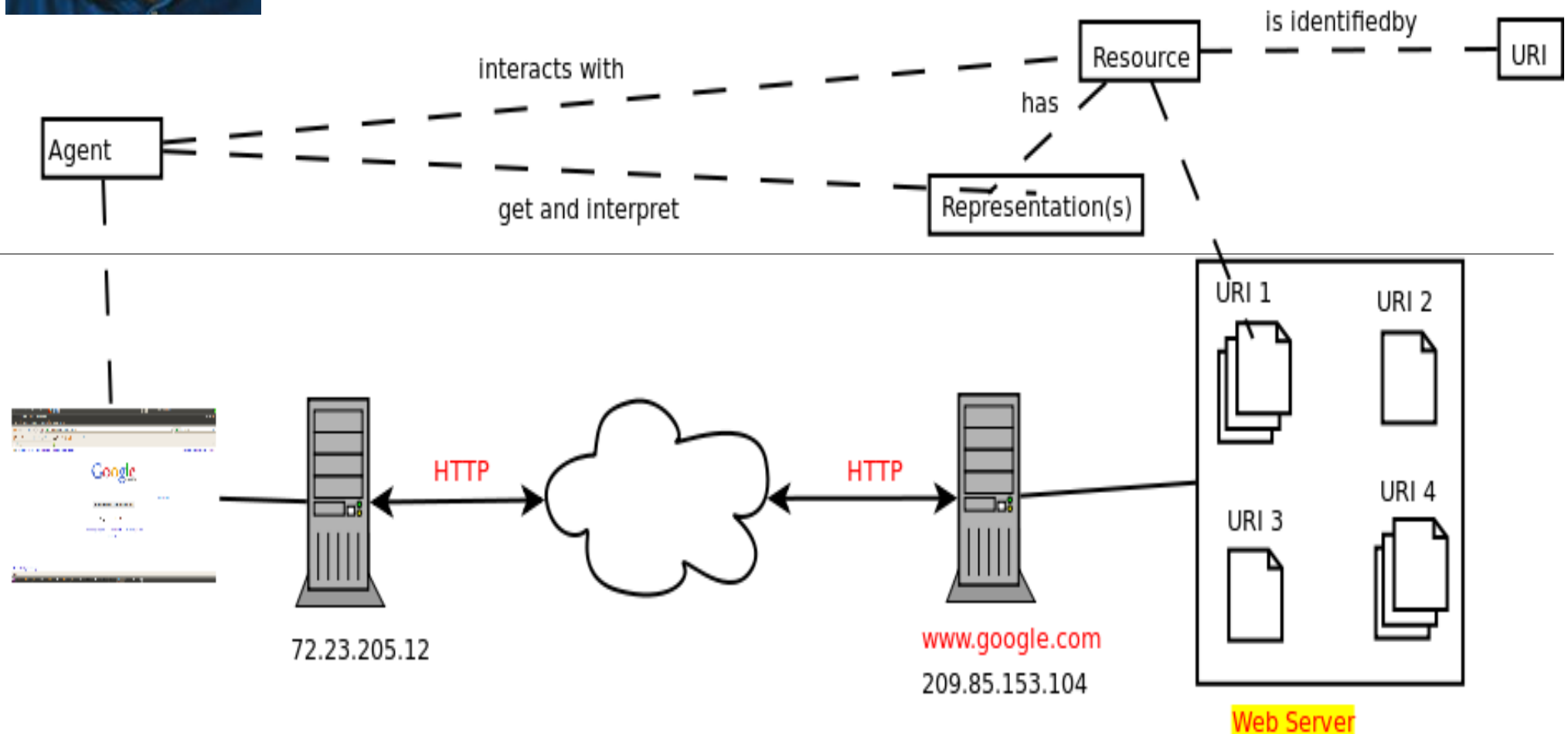
Farooq Ahmed, FAST-NU, Lahore

Farooq Ahmed, FAST-NU, Lahore
Netscape Navigator by Netscape Communications (source : Wikipedia)

Information Architecture of WWW
Representational State Transfer (REST)

Roy Fielding proposed REST in 2000
Picture source: UC Irvine

Summary of WWW Architecture

- WWW is a collection of resources in a web server
- Each resource is identified by one or more URIs which are unique to the resource
- Each resource can have multiple representations e.g. HTML, XHTML, XML, image, etc
- Users use a web agent (e.g. browser, download manager, etc) to access a resource, which is a 3-step process:
  - Identification
  - Interaction
  - Interpretation

Web Development Issues

- Understand the intricacies of web architecture

- Understand technologies (client and server side both) that support efficient creation and consumption of web content for quality user experience

- Build architectures for integration of web-based applications with other aspects of underlying business for supporting diverse functions and allow extensibility in case of changing business needs

- Address performance and scalability issues

- Ensure security of communication over web

REST Phases

- ## Identification

  - Purpose is to locate resources

  - Managed through URIs

  - URI is a general syntax to represent a URL or URN

- ## Interaction

  - Purpose is to retrieve resource representations

  - Managed through HTTP Protocol

  - Other protocols such as HTTPS, SOAP, etc. may also be used

- ## Interpretation

  - Purpose is to understand resource representation and present respective information to user

  - Managed through HTML, CSS, Javascript, etc.

Farooq Ahmed, FAST-NU, Lahore

URI

`http://serc.servehttp.com:8080/wp2011/assignment?id=1#description`

scheme   authority   path   query   fragment

- **Scheme*:** identifies syntax & semantics for URIs using that scheme e.g. http tells agent to use http protocol

- **Authority:** identifies the server address and port. It may also contain optional user info followed by "@" symbol.

- **Path*:** represents the path of resource on the server.

- **Query:** additional information to identify the resource in conjunction with Path

- **Fragment:** represents a sub-resource within the given resource identified by Path & Query information e.g. a para within a page

Note: Constituents marked with * are mandatory, rest are optional and depends upon the scheme and constraints defined in URI definition (see RFC 3986
http://tools.ietf.org/html/rfc3986)

Other URI Examples

- ftp://ftp.is.co.za/rfc/rfc1808.txt

- http://www.ietf.org/rfc/rfc2396.txt

- mailto:John.Doe@example.com

- news:comp.infosystems.www.servers.unix

- tel:+1-816-555-1212

- telnet://192.0.2.16:80/

- urn:oasis:names:specification:docbook:dtd:xml:4.1.2

Communication Protocol

- Set of rules for exchanging messages

    - Data Formats

    - Address Formats

    - Routing

    - Acknowledgements

- Example: Postal Mail

Farooq Ahmed, FAST-NU, Lahore

HTTP Protocol

- HTTP is:

  - Message-based

  - Connection-oriented

  - Request-driven

  - Application-level

  - State-less

HTTP Message

- Message is either:
  - **Request:** from client to server
  - **Response:** from server to client
- Each Message comprises of
  - **Start Line:** defines the message type and other identifying information
  - **Message Header:** contains meta data
  - **Message Body:** contains the actual body of message – generally the entity associated with the message. It may or may not be empty depending upon the Message Type

Entity

- An entity is the content that is intended to be transferred between the client and the server

- Entity may be further divided into:

  - **Header:** That contains meta-information about the body that helps in proper interpretation of body

  - **Body:** The actual content to be delivered

- Examples of Entity:

  - HTML Document

  - XML Document, etc

Typical Sequence of Operations

Server-side technologies:
Servlets/JSP/PHP/ASP.NET etc.

May be used to generate response dynamically

1. client requests a resource

URI 1   URI 2

2. server sends a response (may attach an entity)

URI 3   URI 4

3. client interprets the response

Client side technologies used to interpret the response

Html: basic language
CSS: styling and layout
Javascript: DOM manipulation

Web Server

Request Message

- Request Message is identified by a Request Line, which is of the form:

    Method <SP> Request-URI <SP> HTTP-Version <CRLF>

- Example:

    GET http://www.google.com/index.html HTTP/1.1

    OR

    GET /index.html HTTP/1.1

    Host: www.google.com

Request Methods

GET	Retrieve whatever information is identified by the Request URI
HEAD	Same as get but only fetches the header information of the Entity (e.g HTML Document) and not the actual body itself
POST	The POST method is used to request that the origin server accept the entity enclosed in the request as a new subordinate of the resource identified by the Request-URI in the Request-Line. POST is designed to allow a uniform method to cover the following functions: - Posting a message to a bulletin board, newsgroup, mailing list or similar group of articles; - Providing a block of data, such as the result of submitting a form, to a data-handling process; - Extending a database through an append operation.
PUT	Replace the existing resource with the Entity attached to the request message
DELETE	Delete the resource from the server

Example POST Request

POST http://nu.edu.pk/teacher/add HTTP/1.1

Content-Type: application/x-www-form-urlencoded

Content-Length: 40

name=Farooq Ahmed&course=Web Programming

Message Header

Message Body

Farooq Ahmed, FAST-NU, Lahore

HTTP Response Message

- Response Message is identified by a Status Line of the form:

  HTTP-Version <SP> Status-Code <SP> Reason-Phrase <CRLF>

- Example:

  HTTP/1.1 200   *(means OK: resource found, attached with message)*

  HTTP/1.1 404   *(means Not Found)*

Example Request/Response Message Cycle

- GET http://www.google.com HTTP/1.1

  Host: www.google.com

  Request

- HTTP/1.1 200

  Content-type: text/html

  Content-length: *xx*

  <html>

     <head> … </head>

     <body> Google </body>

  </html>

  Response

HTML

- A markup language
    - Annotates text to add value
    - Uses defined tags with defined semantics

- Lingua-franca of www
- Used to hypertext (link) documents
- Latest version: 5

Farooq Ahmed, FAST-NU, Lahore

Document Structure

- HTML Document comprises of elements
- An element contains other elements
- Everything is contained in HTML element
- HTML element contains:
  - HEAD element: contains meta-information
  - BODY element: contains the actual content

Element Structure

- Element comprises of:
    - Start Tag
    - End Tag
    - Attributes and values
    - Content

<tag attr1="value1" attr2="value2" ...> content... </tag>

<tag attr1="value1" attr2="value2" … />

Basic Elements

Element	Description
<html>	Root element of every html document
<head>	Represents header of document. Contains non-visual elements generally
<body>	Represents the actual content of document
<a>	Used for hyper-linking
<p>	Represents paragraph
 	Represents line-break

Farooq Ahmed, FAST-NU, Lahore

Muhammad Iqbal

Sir Muhammad Iqbal (November 9, 1877 – April 21, 1938), also known as Allama Iqbal, was a philosopher, poet and politician[1] in British India who is widely regarded as having inspired the Pakistan Movement. He is considered one of the most important figures in Urdu literature,[2] with literary work in both the Urdu and Persian languages.[1][2]
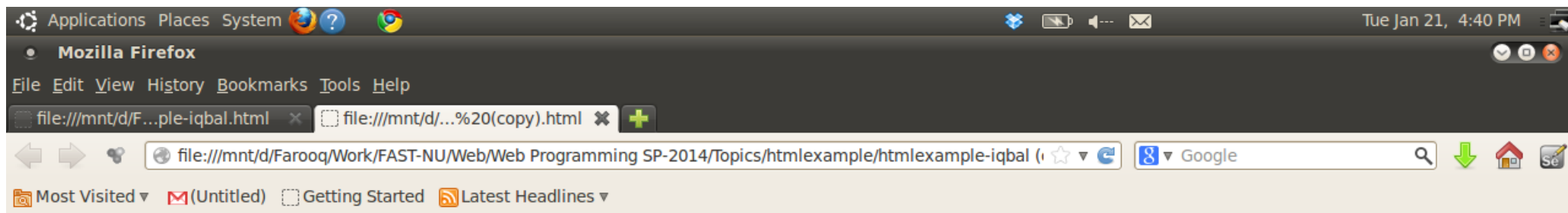
Iqbal is admired as a prominent classical poet by Pakistani, Indian, Iranian, and other international scholars of literature.[3][4] Though Iqbal is best known as an eminent poet, he is also a highly acclaimed "Muslim philosophical thinker of modern times".[1][4] His first poetry book, Asrar-e-Khudi, appeared in the Persian language in 1915, and other books of poetry include Rumuz-i-Bekhudi, Payam-i-Mashriq and Zabur-i-Ajam. Amongst these his best known Urdu works are Bang-i-Dara, Bal-i-Jibril, Zarb-i Kalim and a part of Armughan-e-Hijaz.[5] In Iran and Afghanistan, he is famous as Iqbal-e Lahore (Iqbal of Lahore), and he is most appreciated for his Persian work.[6] Along with his Urdu and Persian poetry, his various Urdu and English lectures and letters have been very influential in cultural, social, religious and political disputes over the years

…

Excerpt on Iqbal taken from Wikipedia

Farooq Ahmed, FAST-NU, Lahore

Mozilla Firefox

File Edit View History Bookmarks Tools Help

file:///mnt/d/Fa...ample-iqbal.html

file:///mnt/d/Farooq/Work/FAST-NU/Web/Web Programming SP-2014/Topics/htmlexample/htmlexample-iqbal.h   Google

Most Visited ▾   (Untitled)   Getting Started   Latest Headlines ▾

Muhammad Iqbal Sir Muhammad Iqbal (November 9, 1877 - April 21, 1938), also known as Allama Iqbal, was a philosopher, poet and politician[1] in British India who is widely regarded as having inspired the Pakistan Movement. He is considered one of the most important figures in Urdu literature,[2] with literary work in both the Urdu and Persian languages.[1][2] Iqbal is admired as a prominent classical poet by Pakistani, Indian, Iranian, and other international scholars of literature.[3][4] Though Iqbal is best known as an eminent poet, he is also a highly acclaimed "Muslim philosophical thinker of modern times".[1][4] His first poetry book, Asrar-e-Khudi, appeared in the Persian language in 1915, and other books of poetry include Rumuz-i-Bekhudi, Payam-i-Mashriq and Zabur-i-Ajam. Amongst these his best known Urdu works are Bang-i-Dara, Bal-i-Jibril, Zarb-i Kalim and a part of Armughan-e-Hijaz.[5] In Iran and Afghanistan, he is famous as Iqbal-e-Lahori (Iqbal of Lahore), and he is most appreciated for his Persian work.[6] Along with his Urdu and Persian poetry, his various Urdu and English lectures and letters have been very influential in cultural, social, religious and political disputes over the years ...

Rendering of text by Firefox browser
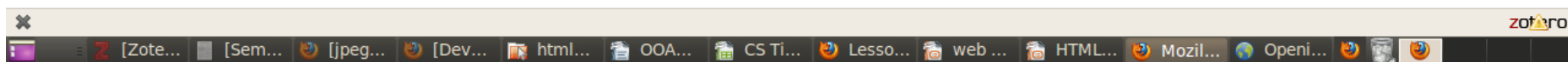
```html
<html>
<body>
<h1>Muhammad Iqbal</h1>

<p>
Sir Muhammad Iqbal (November 9, 1877 - April 21, 1938), also known as Allama Iqbal, was a philosopher, poet and politician[1] in British India who is widely regarded as having inspired the Pakistan Movement. He is considered one of the most important figures in Urdu literature,[2] with literary work in both the Urdu and Persian languages.[1][2]
</p>
<p>
Iqbal is admired as a prominent classical poet by Pakistani, Indian, Iranian, and other international scholars of literature.[3][4] Though Iqbal is best known as an eminent poet, he is also a highly acclaimed "Muslim philosophical thinker of modern times".[1][4] His first poetry book, Asrar-e-Khudi, appeared in the Persian language in 1915, and other books of poetry include Rumuz-i-Bekhudi, Payam-i-Mashriq and Zabur-i-Ajam. Amongst these his best known Urdu works are Bang-i-Dara, Bal-i-Jibril, Zarb-i Kalim and a part of Armughan-e-Hijaz.[5] In Iran and Afghanistan, he is famous as Iqbal-e-Lahori (Iqbal of Lahore), and he is most appreciated for his Persian work.[6] Along with his Urdu and Persian poetry, his various Urdu and English lectures and letters have been very influential in cultural, social, religious and political disputes over the years
</p>

<p>...</p>
</body>
</html>
```
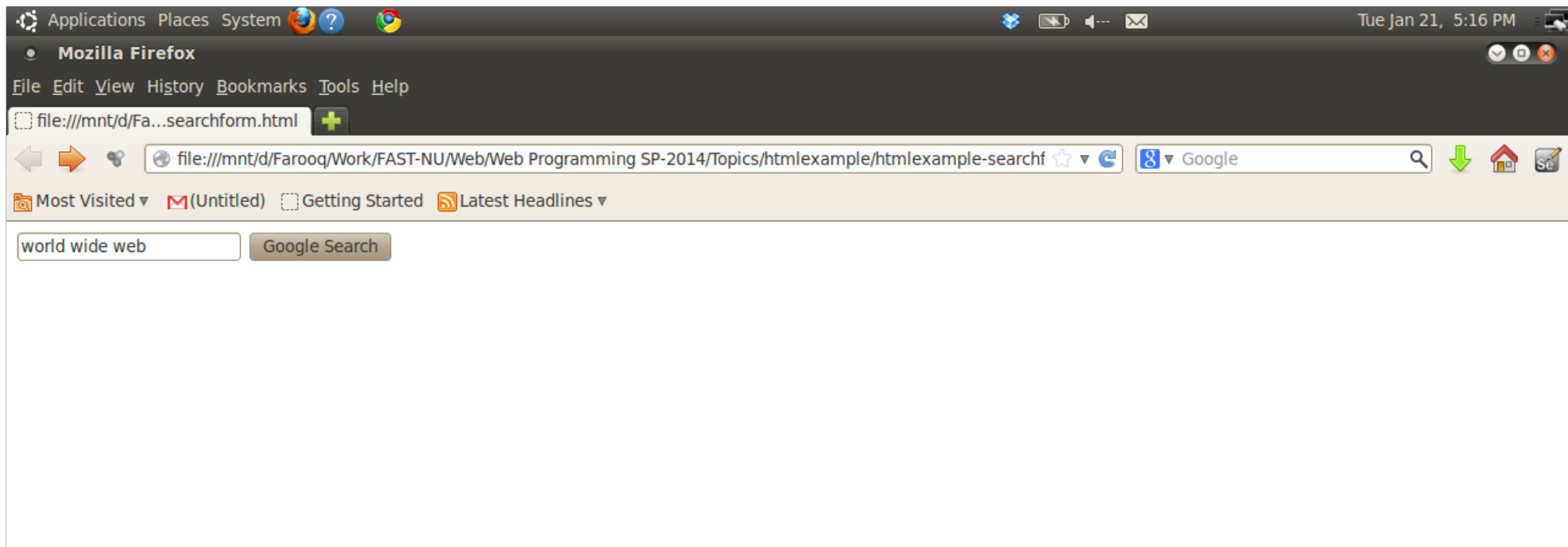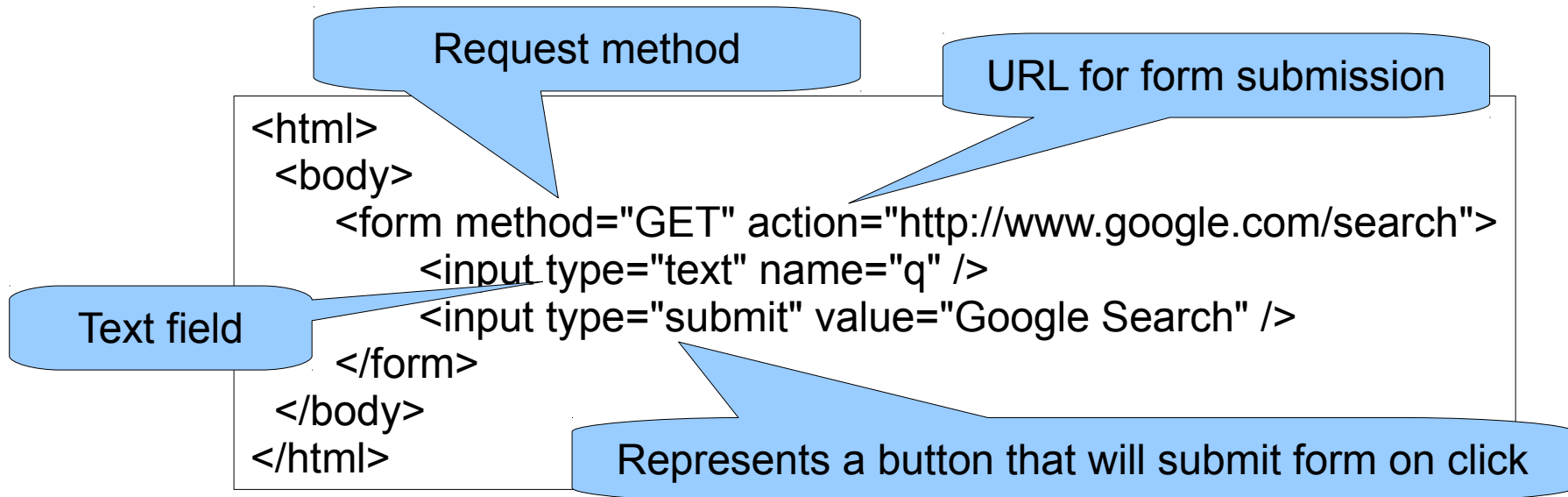
Excerpt annotated using html

Muhammad Iqbal

Sir Muhammad Iqbal (November 9, 1877 - April 21, 1938), also known as Allama Iqbal, was a philosopher, poet and politician[1] in British India who is widely regarded as having inspired the Pakistan Movement. He is considered one of the most important figures in Urdu literature,[2] with literary work in both the Urdu and Persian languages.[1][2]

Iqbal is admired as a prominent classical poet by Pakistani, Indian, Iranian, and other international scholars of literature.[3][4] Though Iqbal is best known as an eminent poet, he is also a highly acclaimed "Muslim philosophical thinker of modern times".[1][4] His first poetry book, Asrar-e-Khudi, appeared in the Persian language in 1915, and other books of poetry include Rumuz-i-Bekhudi, Payam-i-Mashriq and Zabur-i-Ajam. Amongst these his best known Urdu works are Bang-i-Dara, Bal-i-Jibril, Zarb-i Kalim and a part of Armughan-e-Hijaz.[5] In Iran and Afghanistan, he is famous as Iqbal-e-Lahori (Iqbal of Lahore), and he is most appreciated for his Persian work.[6] Along with his Urdu and Persian poetry, his various Urdu and English lectures and letters have been very influential in cultural, social, religious and political disputes over the years

...

Rendering of html by Firefox browser

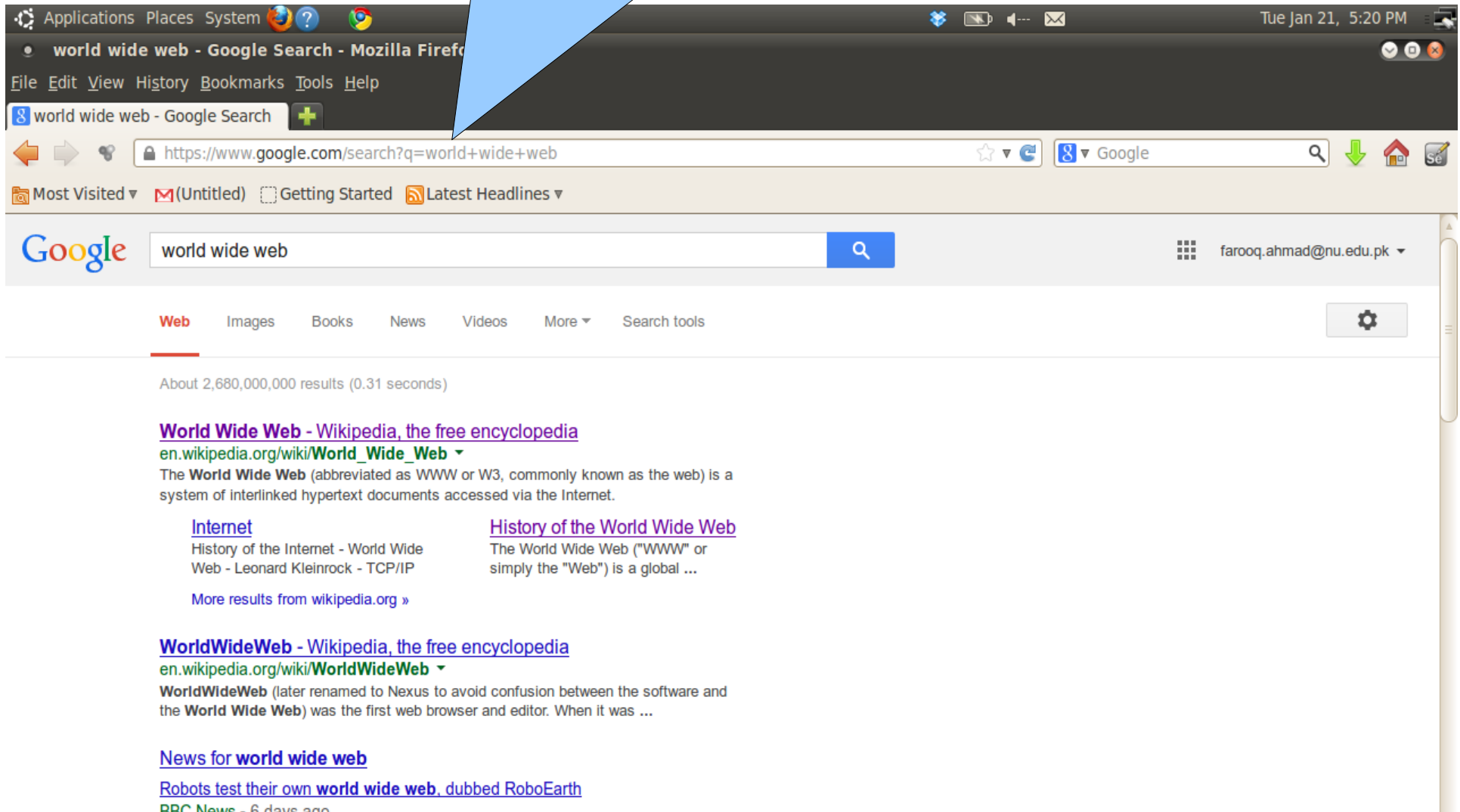Farooq Ahmed, FAST-NU, Lahore

Form Elements

Element	Description
<form>	Represents a form element
<fieldset>	Used to logically group together form elements
<label>	Defines a label for an input element
<input>	Specifies a field to take input from user
<textarea>	Represents a larger area for textual input instead of just a field
<select>	A drop-down list
<option>	Represents an option within select list
<button>	Represents a button

Types of Input

Type	Description
button	Represents a form button
checkbox	Represents a checkbox
file	Represents a file dialog
hidden	Represents a hidden input field – hidden from user
password	Represents a password field
radio	Represents a radio button
reset	Represents a button that reset form inputs to default
submit	Represents a button that submits form input to action URL
text	Represents a text input field

```html
<html>
  <body>
    <form method="GET" action="http://www.google.com/search">
      <input type="text" name="q" />
      <input type="submit" value="Google Search" />
    </form>
  </body>
</html>
```

Request method

URL for form submission

Text field

Represents a button that will submit form on click

Form parameters are shown in browser bar

Farooq Ahmed, FAST-NU, Lahore

Request method

URL for form submission

```html
<html>
<body>
    <form method="POST" action="http://slate.nu.edu.pk/portal/xlogin">
        User <input type="text" name="eid" />
        Password <input type="password" name="pw" />
        <input type="submit" value="login">
    </form>
</body>
</html>
```

Text field

Password field

Applications   Places   System   Tue Jan 21, 8:20 PM

Mozilla Firefox

File Edit View History Bookmarks Tools Help

Inbox (1,322) - farooq.ahma...   ✕   file:///mnt/d/F...form-post.html   ✕   ✚

file:///mnt/d/Farooq/Work/FAST-NU/Web/Web Programming SP-2014/Topics/htmlexample/htmlexample-loginfor   ▼   Google

Most Visited ▼   M (Untitled)   Getting Started   Latest Headlines ▼

User farooq.ahmad   Password ●●●●●●●●   login

Farooq Ahmed, FAST-NU, Lahore

Form parameters not shown

GET vs POST

- GET requests can be cached

- GET requests remain in the browser history

- GET requests can be bookmarked

- GET requests have length restrictions

- GET requests are to retrieve data

- POST requests are never cached

- POST requests do not remain in the browser history

- POST requests cannot be bookmarked

- POST requests have no restrictions on data length

- POST requests should be used to transmit data

GET requests should never be used when dealing with sensitive data

HTML5

- Elements
    - Semantic: header, footer, section, navigation, etc.
    - Form input: number, date, time, calendar, range, etc.
    - Graphics: svg, canvas
    - Multimedia: audio, video
- Related APIs (published separately)
    - Web workers
    - Web Storage
    - Geo-location
    - Many others …