

Session 2

Background

1

Lecture Objectives

- Understand how an Internet resource is accessed
- Understand the high level structure of the "Internet cloud"
- Understand the high level structure of the TCP/IP protocols
- Understand how a computer finds the IP address of a host using DNS
- Know the structure of MIME type standards

© Robert Kelly, 2001-2012

Reference

- Mime types

http://www.w3schools.com/media/media_mimeref.asp

- Cyber Security

- Cyber War by Richard Clarke and Robert Knake, Harper Collins, 2010

© Robert Kelly, 2001-2012

3

What is a URL?

- A short string that identifies a resource on the Web
- Stands for Uniform Resource Locator
 - Uniform - varied and new types of resources
 - Resource - Anything that has identity (e.g., image)
- Reduces the tedium of connecting to a host, selecting a path, selecting a resource, etc. into a single string that:
 - Can be saved as a bookmark in your browser
 - Can be saved as an object in your Java code

© Robert Kelly, 2001-2012

URL Example

http://www.cs.sunysb.edu

Protocol identifier *Resource name*

- Protocol Identifier - Indicates the name of the protocol to be used to fetch the resource
- Resource name is the complete address of the resource
- Resource name may be appended with a fragment / reference / named anchor (denoted by #) or include a query string (denoted by ?)

© Robert Kelly, 2001-2012

URL Resource Name

- For http, the name includes: host name, path name to the file, port number (optional), and location within the resource (optional)
- The resource is not necessarily a file, it can be generated on the fly ← Very important concept
- A trailing slash (www.sun.com/) is shorthand for the file named /index.html

© Robert Kelly, 2001-2012

Relative URLs

- Contains only enough information to reach the resource relative to (in the context of) another URL
- Used within HTML files
- specified relative to the location of the file in which they are contained

```
<td> <a href="LectureCode.html" >Intro</a></td>
```

URL context =

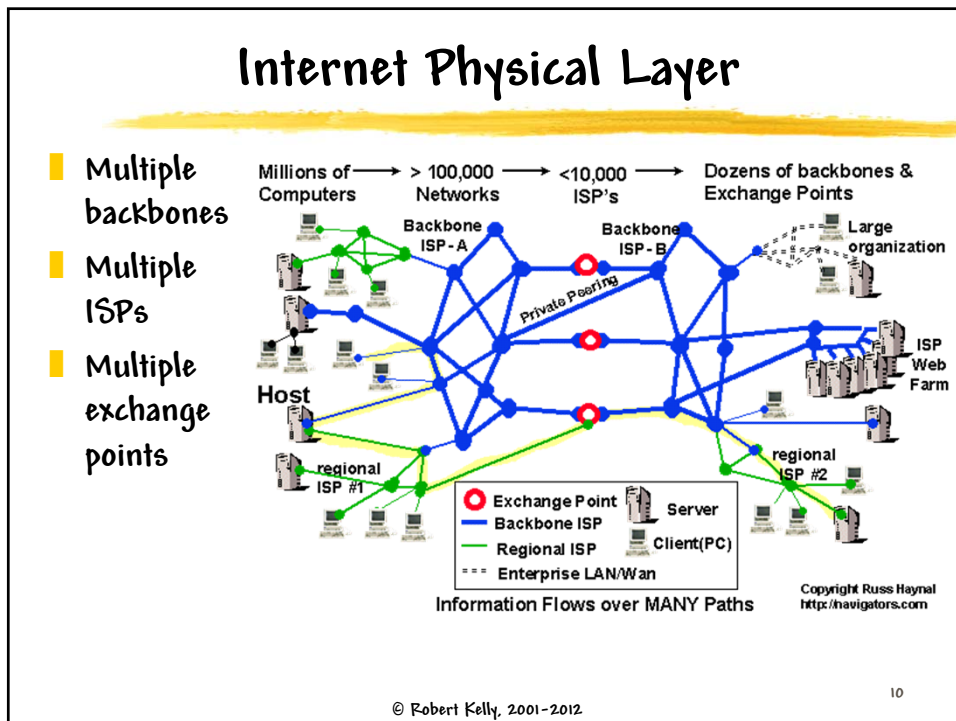
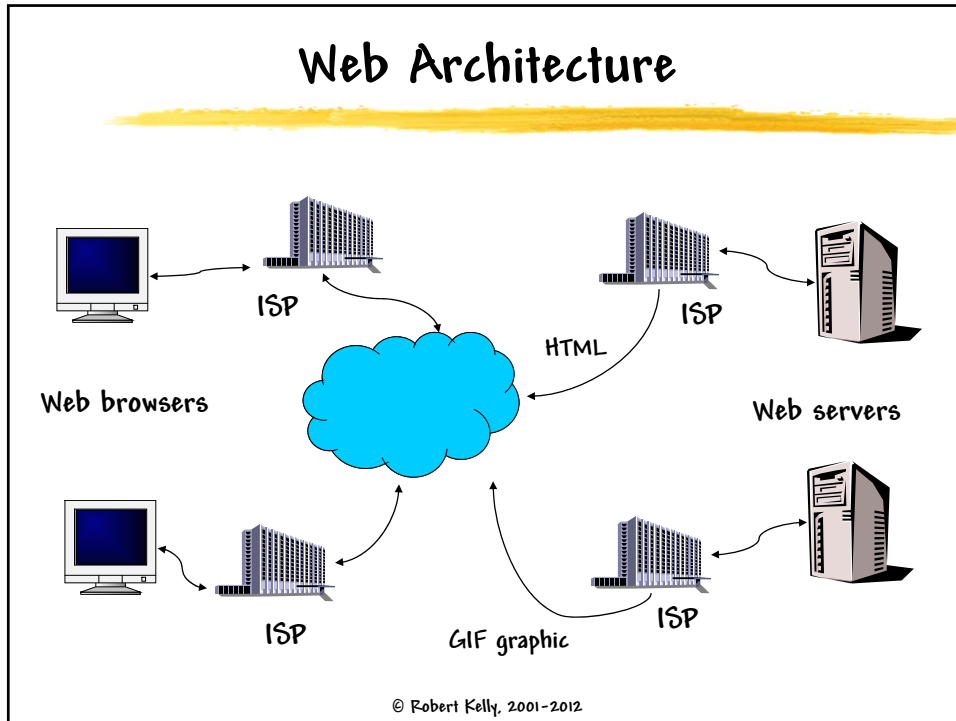
```
new URL("http://www.sciam.com/feature_directory.cfm");  
URL aURL = new URL(context, "myImage.gif");
```

© Robert Kelly, 2001-2012

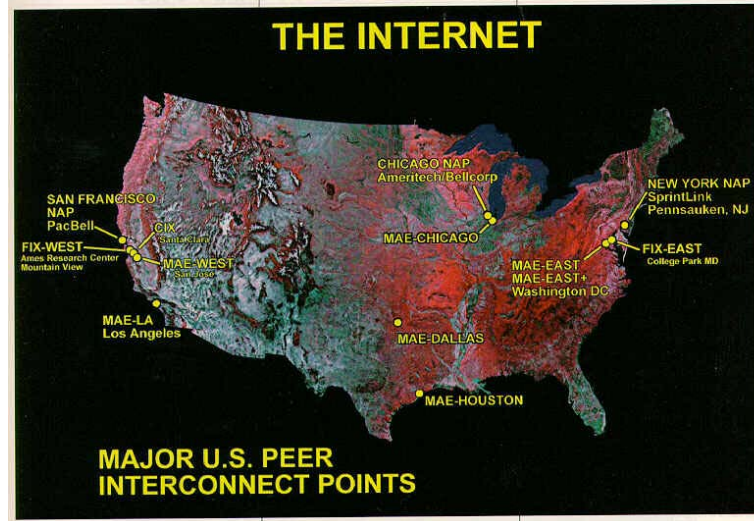
Examples of URL Schemes

- The standard Java library currently provides support for http and ftp
- Other protocols can be used with some extra coding
 - tel:+358-555-1234567
 - fax:+358.555.1234567
 - modem:+3585551234567;type=v32b?7el;type=v110
 - tel:+358-555-1234567;postd=pp22
 - mailto:majordomo@example.com?body=subscribe%20bamboo-l

© Robert Kelly, 2001-2012



Internet Interconnect Points



© Robert Kelly, 2001-2012

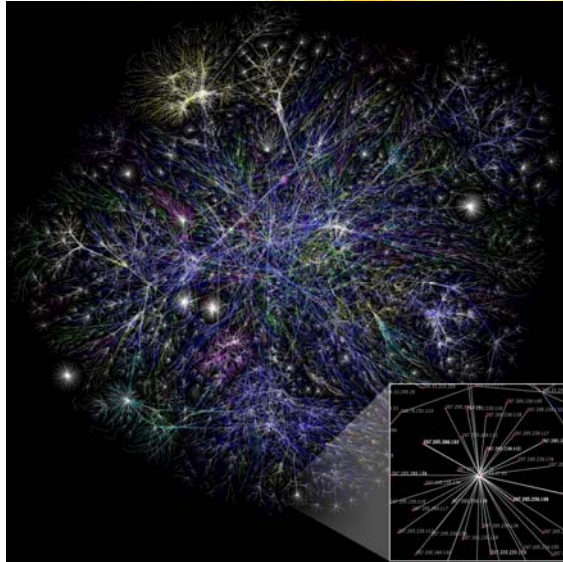
US Backbone Providers

- Tier 1 ISPs
 - Verizon
 - AT&T
 - Qwest
 - Sprint
 - Level 3
 - Global Crossing

© Robert Kelly, 2001-2012

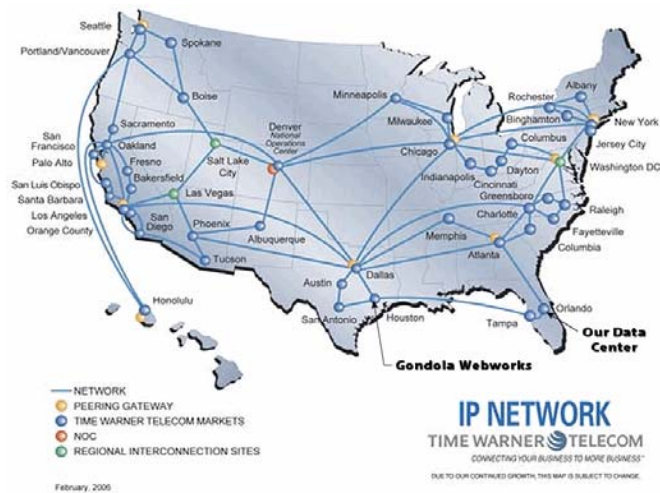
12

Internet

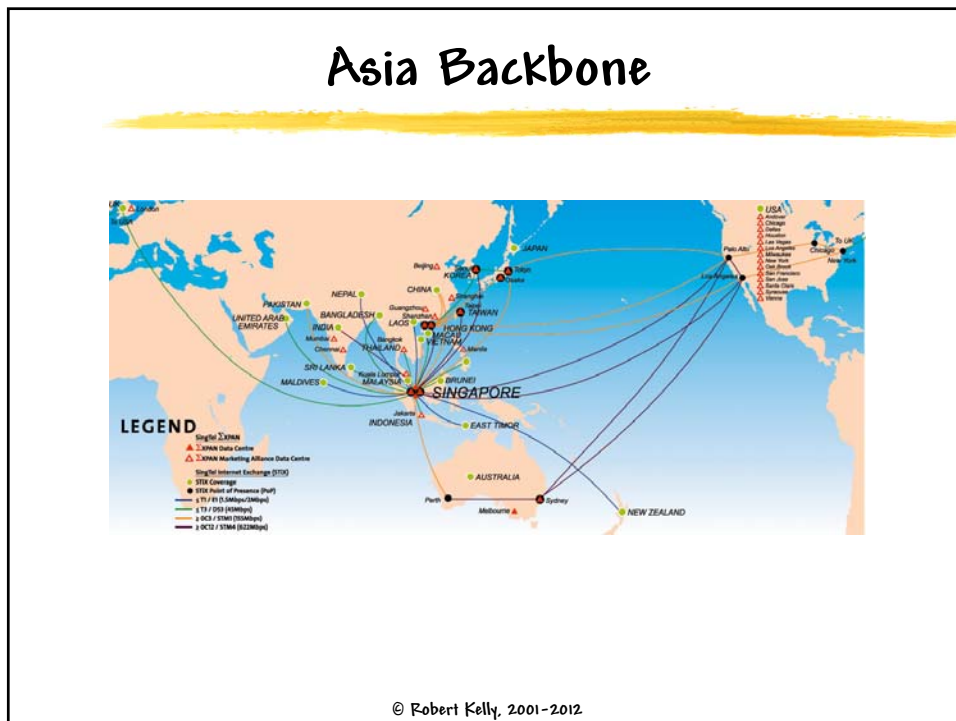
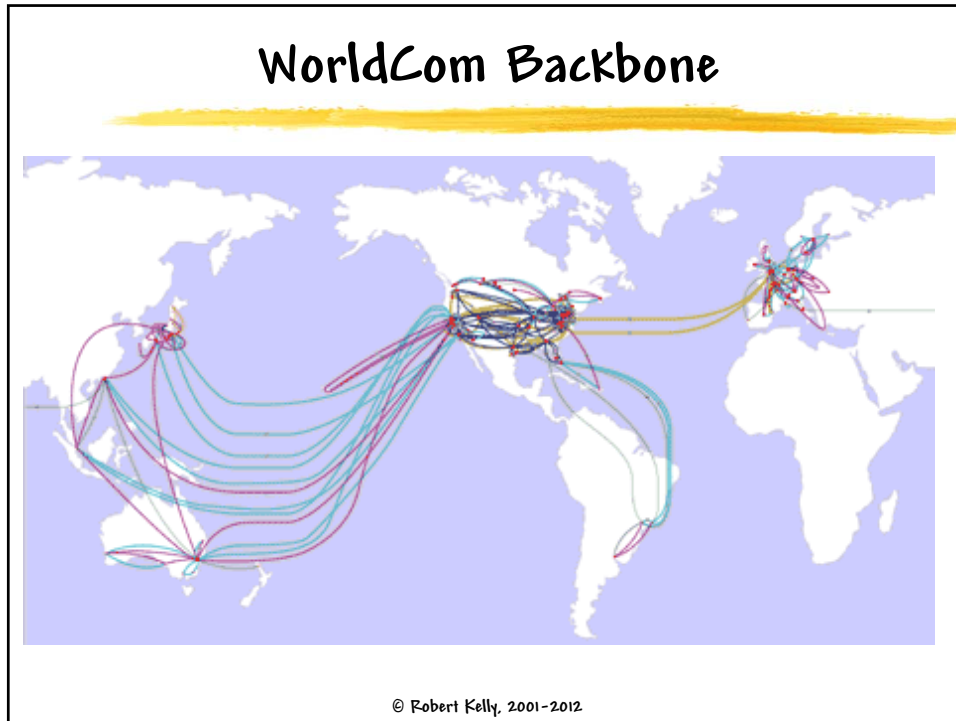


© Robert Kelly, 2001-2012

Time Warner Network Map



© Robert Kelly, 2001-2012



Terms

- Protocols - control the sending and receiving of information over the Internet
- Physical media - cable, copper wire, fiber, radio spectrum
- Routers - intermediate switching devices
- Route (or path) - sequence of physical media and routers to complete end-to-end communication
- Packet switching - decomposing a message into packets and routing the packets to a destination

© Robert Kelly, 2001-2012

Transport Protocols

- Internet connection-oriented service is implemented through buffers at the sending and receiving end

TCP

- Transmission Control Protocol
- Connection based
- Reliable flow of data between two computers

UDP

- User Datagram Protocol
- Connectionless service
- Order of delivery is not guaranteed

© Robert Kelly, 2001-2012

TCP/IP

- IP (Internet Protocol) - network layer protocol that specifies the format of information that is sent and received among routers and end nodes

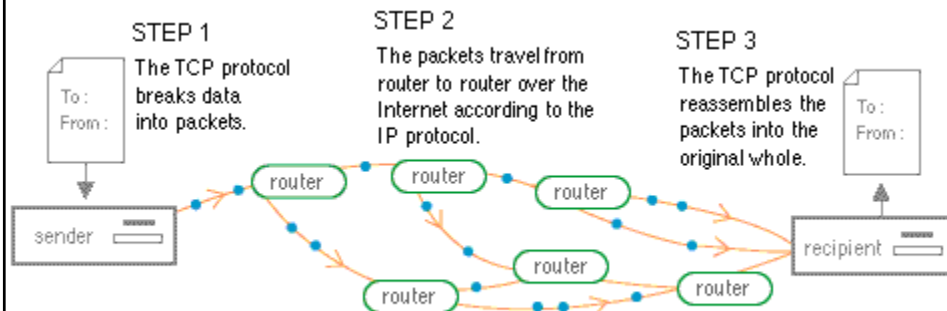
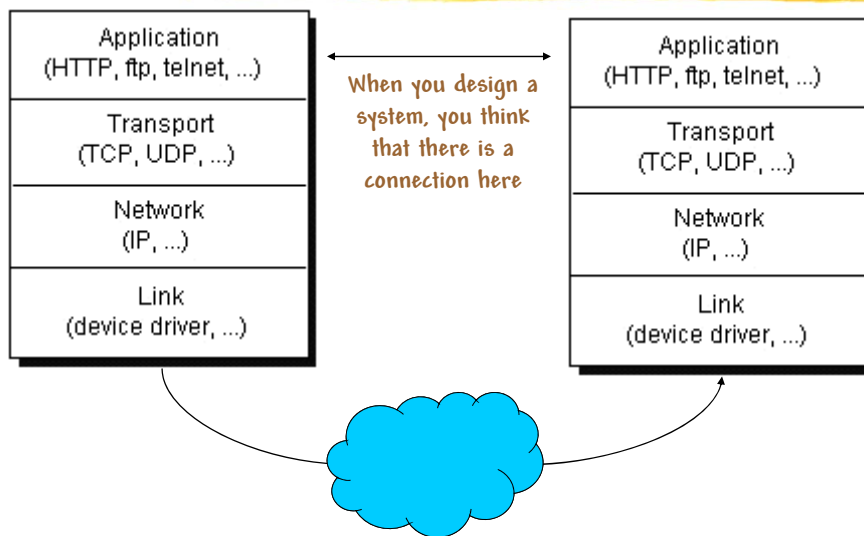


Figure 2. How data travels over the Net.

© Robert Kelly, 2001-2012

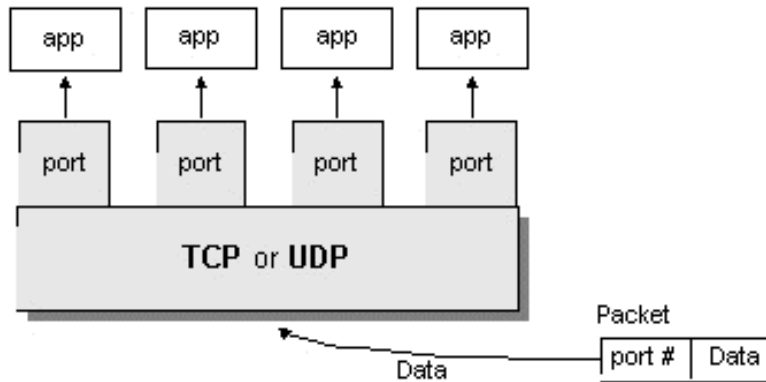
Internet Protocol Stacks



© Robert Kelly, 2001-2012

Port

- To communicate with a computer, you need to specify the address of the host computer and the identity of the receiving process on the host computer



© Robert Kelly, 2001-2012

Typical Port Number Assignments

Service	Port
telnet	23
smtp	25
whois++	63
http	80

© Robert Kelly, 2001-2012

22

How Does a Browser Work?

- It reads a named resource (usually an HTML document) on an Internet-based server
- Begins to display the page
- Finds all the URLs in the HTML
- Requests the resources associated with the other URLs (e.g., images)
- Includes the additional resources in the display of the page

© Robert Kelly, 2001-2012

23

How Does the Internet Find a Host?

- URL contains a host name, (e.g., `www.cnn.com`)
- Internet routers identify hosts by their IP address
 - 4 bytes, presented in dotted-decimal notation
 - decimal numbers, separated by periods (e.g., `121.7.106.83`)
 - Each number is between 0 and 255
 - The 4 fields identify 4 levels in the network hierarchy
 - 2^{32} possible IP addresses (in IP V4)
- DNS (Domain Name System) is a directory service that translates host names (sometimes referred to as domain names) into IP addresses

© Robert Kelly, 2001-2012

24

How DNS Works

The diagram illustrates the DNS lookup process. A browser sends a request to a DNS client. The DNS client then queries a local name server. If the local server doesn't have the answer, it queries the root name server. The root server then directs the query to the appropriate authoritative name server.

- Browser passes host name to DNS client.
- DNS client looks up IP address from the distributed database located on the DNS name servers
 - Local
 - Root
 - Authoritative

25

© Robert Kelly, 2001-2012

Domain Name Service

- Used to map host names (contained within a URL) into network addresses (32 bit IP address)
- Name space (partial)

```
graph TD; root["(root)"] --- mil; root --- gov; root --- net; root --- com; root --- edu; root --- org; root --- dots["..."]; com --- us; com --- them; us --- another;
```

26

© Robert Kelly, 2001-2012

Other DNS Services

- Host aliasing - a host can have one or more alias names (e.g., `ibm.com` and `www.ibm.com`), one of which is the canonical hostname
 - Aliasing by service - a company can use the same host name for its Web server, mail server, ftp server, etc. even though these are different computers with different IP addresses
 - Load distribution - a set of IP addresses can be associated with a canonical hostname. DNS will return the set of addresses, varying the order of addresses each time it is accessed
- How many hosts do you think are associated with `google.com`?

© Robert Kelly, 2001-2012

27

DNS Concepts to Remember

- DNS provides name to address mapping
- DNS is implemented through local address caching
- DNS is a distributed database service
 - Very reliable
 - Not always fast
 - Not always up to date

© Robert Kelly, 2001-2012

28

Cyber Security Issues

- Internet is inherently insecure
- Industries and country economies are increasingly dependent on the Internet
- Many successful cyber attacks
- Examples of cyber warfare (e.g., Estonia and Georgia)
- Development of cyber warfare capabilities by some countries

© Robert Kelly, 2001-2012

29

Vulnerabilities of the Internet¹

- Routing among ISPs - Border Gateway Protocol routes packets across the Internet, but there is no checking of authenticity of messages
- No governance of the Internet beyond domain names *An advanced packet sniffer on an Ethernet network can look at all the traffic*
- Operational messages are unencrypted
- Malware can be easily propagated
- Decentralized design

1. Clarke, R. Cyber War

© Robert Kelly, 2001-2012

30

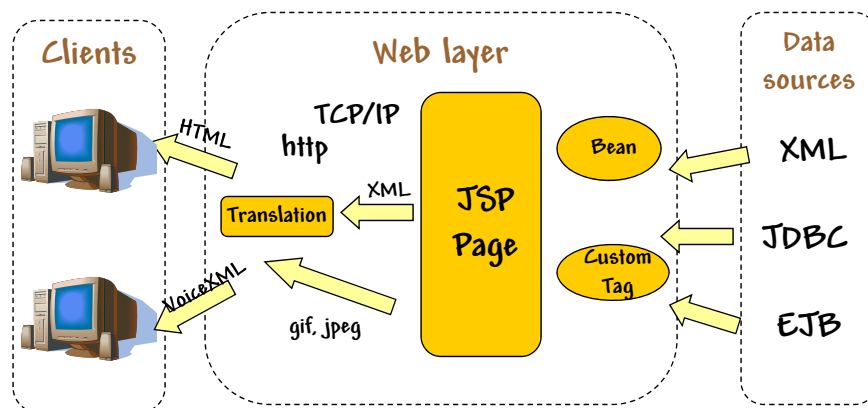
What is a MIME Type?

- Multipurpose Internet Mail Extensions
- Designed for the interchange of data among various e-mail systems
- Allows for universal interchange of data
- Defines naming of file types
- Organized into 8 base type categories

© Robert Kelly, 2001-2012

31

Extended Web Architecture



The host-host data transfers usually involve one MIME file

© Robert Kelly, 2001-2012

32

"Base" Media Types

- application - application and miscellaneous
- audio - audio data
- image - image data e.g., image/png
- message - news, e-mail, etc.
- model - models (e.g., geometric)
- multipart - multipart
- text - HTML, CSS, etc.
- video - video data
- unregistered - name usually begins with an "x-"
 - x-troff-ms

© Robert Kelly, 2001-2012

33

MIME Type Concepts to Remember

- Internet files have standard formats so that data can be exchanged easily between very different computers (hardware, OS, etc.)
- Sender specifies the type and receiver interprets the data accordingly, taking into account all the differences in internal data representation

© Robert Kelly, 2001-2012

34

Did You Achieve the Lecture Objectives?

- Understand how an Internet resource is accessed
- Understand the high level structure of the "Internet cloud"
- Understand the high level structure of the TCP/IP protocols
- Understand how a computer finds the IP address of a host using DNS
- Know the structure of MIME type standards