B211 Internet Computing

Internals of the Web Client

Learning Objectives

1. Understand the basic components and operations of a web client.

2. Understand the relationship between what the web client does, and what the web server does.

Lecture Outline

- The role of the web client
- Common features of a web browser
- Current web browser usage

The Web Client

- Web Clients are usually browsers like Netscape Navigator and Microsoft Internet Explorer.
    - But they need not be so. Any software written to the specification of HTTP, and communicates with a web server is a web client.
- The reason why most web clients are (what we know as) browsers is because most resources transferred through HTTP are web pages with multimedia, which requires and interface such as that of browsers to effectively view.

The Universal Interface

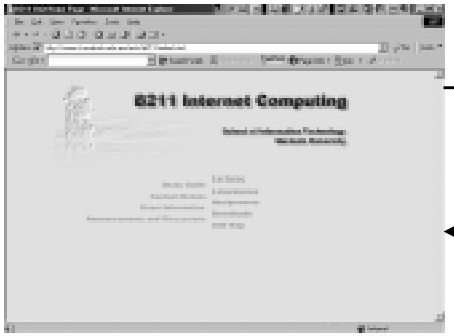
- Browsers are meant to act as a universal interface to all web-based communication, or a universal client which handles functions previously handled by other clients.
- For most novice users today introduced to a web-capable machine, the browser effectively takes over most of the functions of an operating system interface.

Format Negotiation

- HTTP allows for format negotiation. The feature of is to allow the browsers and server to transmit any form of data formats which they may conceivably want to transmit.
- The client negotiates with the server (using the *Accept* field in a HTTP request), what format data they can handle.

Basic Operations of a Web Client

1. User types in URL www.it.murdoch.edu.au/units/b211



2. Browser assumes it is requesting through HTTP, gets the hostname and resource path from the URL.

GET /units/b211/ HTTP/1.1
...

3. Send request for resource to hostname.

HTTP/1.1 200 OK
Content-type: text/html
…
<html>
…
</html>

4. Get response .

5. Process response and display if appropriate.

Some Common Features of Browsers

- Some common features of most (not all) browsers:
  - MIME Helper Applications
  - Cookies
  - Push Technologies
  - Plug-ins
  - Applets

Plug-ins and Helper Applications

- Browsers understands and can display at least HTML and some common file formats (eg. GIF and JPG).
- But it can accept ALL possible formats
  - One of the original design features of HTTP was that clients/servers must be able to exchange ANY format files.
  - MIME messages used by HTTP can carry any format data.
- For formats it does not understand, it uses a local helper or plug-in application to decipher (and perhaps display) it.

Plug-ins and Helper Applications

- The list of which application is responsible for which file format is kept by the browser or (in the case of IE in Windows) the client operating system.
- Common examples:
  - WORD and EXCEL documents in IE and Windows.
  - Adobe Acrobat Reader for PDF files.
  - Macromedia Flash Player for Flash animations
  - Quicktime player for movies

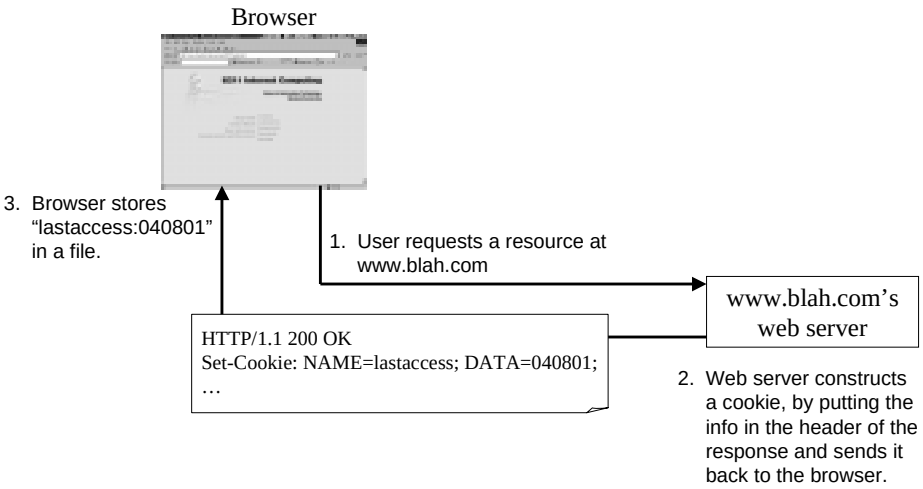
HTTP Cookies

- As mentioned before, each HTTP transaction is independent.
- But servers want to keep information about a client, to be retrieved later.
  - Eg. a users preference, or previous transactions in a electronic shopping, etc.
- Cookies were originally proposed by Netscape, later adopted by other browsers.

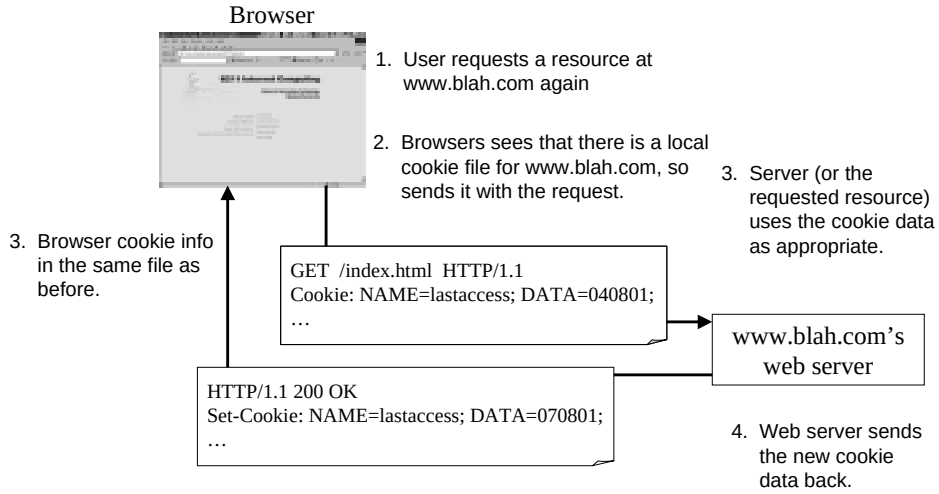
HTTP Cookies

- Cookies are sent by servers to clients (browsers), and are managed and stored locally by the clients. Clients send such cookie data back to the same server at later sessions.
- Clients have the ability to reject such cookies sent by servers.
  - You will see this in next week (week 4's labs).

Example Cookie Transaction (first time)



Example Cookie Transaction (first time)



The Cookie Myth

- A common misconception about cookies is that it allows servers to get information about you and your machine beyond what it is suppose to.
- Servers can only store in cookies data which it already has received through HTTP request, so basically you've already given the server the information.
    - You should be careful what information you actively supply originally, rather than about the cookie.

Push Technology

- Extends the communication between client/servers to more than just a single transaction, to a few transactions where the data is related in some way.
- Conventional web transaction works by client pulling information from a server. Push technology is about having servers push information onto the clients without it being initiated by the clients everytime.
- Examples are such technologies are telecasts of events over the web (Webcasting), news and sports updates, etc.

Push Technology (cont'd)

- Different ways of implementing push technology:
  - By Notification:
    - Client registers with a web server and awaits notification that something interesting has happened.
  - By Scheduled Pull:
    - Client pulls information (such as web pages) at per-determined intervals
  - By Channels
    - Like pay-per-view TV
    - Client subscribes to a certain "streams", and server sends the streams to all the subscribers
    - This method requires support of multicasting to scale properly to large number of subscribers

JAVA Applets

- Applets are applications which are embedded in another application, usually HTML page.
- Instead of having remote applications execute on server-side, applets are downloaded and executed on the client-side, usually in a JAVA-enabled browser (ie. a browser with a JAVA *run-time engine*).
- Security issues becomes very important
  - applets not allowed to access a client's system
  - applets can only talk with the server hosting it

Current Browser Usage

- Estimates on browser usage is a very hard problem.
  - Web servers are a lot easier since they correspond to web-sites, and web-sites are usually publically advertised and available.
- We can only get statistics on browsers by looking at the logs on web servers,
  - But which web servers do we use? Different people, groups and communities access their own very specific set of web sites.

Current Browser Usage

- Some example sites which publishes updated statistics on browsers which accesses the sites:
  - http://browserwatch.internet.com/stats/stats.html
  - http://www.cen.uiuc.edu/~ejk/bryl.html
  - http://www.co.broward.fl.us/usage/brow0501.htm
  - http://wwwseti.cs.utwente.nl/ServerStats/OldBrowsers.html
- You can see from the above links that the figures are very different depending on the surfers on the different sites.

References:

- For a chart of comparisons between the features of different browsers:
  - http://hotwired.lycos.com/webmonkey/reference/browser_chart/
- For up-to-date news on browser developments:
  - http://browserwatch.internet.com/