Authentication & .htaccess

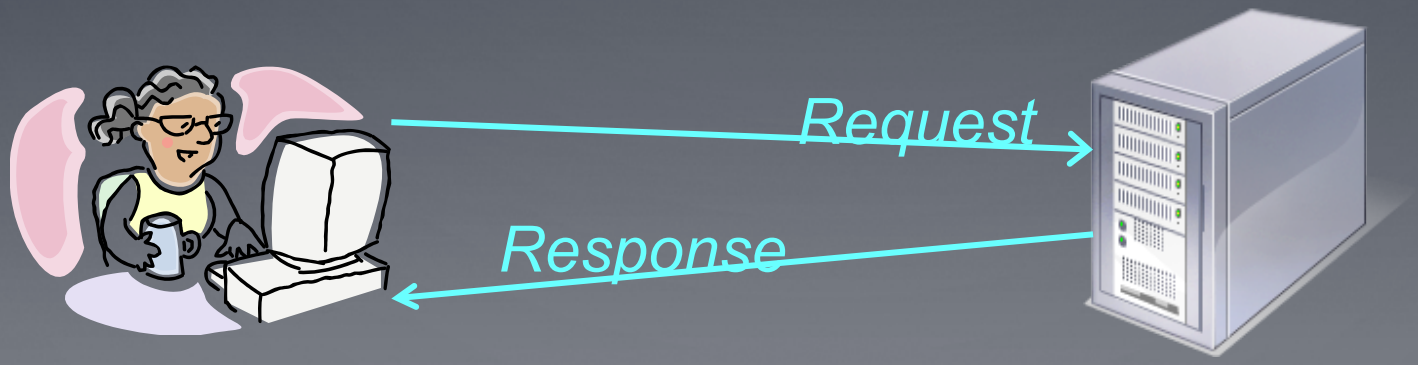
Basic Web authentication and using Apache's .htaccess files

HTTP Servers

- Web pages are hosting by HTTP servers
- Web servers send and receive data using the Hypertext Transfer Protocol (HTTP)
- Servers can send & receive data on virtually any protocol, but HTTP is the standard for web pages

How's it work?

- The Web uses a "request/response" architecture
- The typical process is:
  - The browser sends a *request* to the server
  - The server locates the requested HTML file and reads it
  - The server sends the HTML code to the browser



Server Configurations

- Every Apache server has a master server configuration file
- Apache allows you to make folder (directory) specific configuration files that override the master defaults
- These files allow you to customize your hosting for tasks like authentication, error handling (like 404 pages), redirects, etc.

Server Configurations

- Ideally, you use the master configuration of the server.
- If for some reason you don't have permissions to it (like in the case of Banjo), then you can use .htaccess files.
- Other servers may not have these specific configurations.

  Many instead just use code! ☺

Server Configurations

- Server-side code is the most powerful and flexible, but that means that it must be programmed.
- Configuration files are much faster, but they are not as flexible or powerful.
- .htaccess files can become very difficult to use as the complexity of the site increases. For simple tasks, they are very quick to make.

Apache

- Apache uses configuration files called **.htaccess** files
- These are Hypertext Access files (htaccess)
- These files always start with a dot

  .htaccess

R·I·T

.htaccess

- .htaccess files take commands called directives
- Each directive allows you to specify certain behaviors
- This means that the server has to constantly open and read the .htaccess files, which can be bad for performance.

.htaccess

- .htaccess files are very fast to setup and use for small tasks
- You can only have one .htaccess file per folder.
- Each .htaccess file will affect the folder it is in and the sub-folders under it.

.htaccess

- Each sub-folder could again have its own .htaccess file, but it would hurt performance.
- Every .htaccess file slows down the server
- .htaccess files are case-sensitive

DirectoryIndex

- The DirectoryIndex directive tells Apache which file is the index.

- The default is index.html, but you can provide a different name. The directive takes a file name from the same folder.

  Example of overriding the default index.html

  DirectoryIndex main.html

HTTP Status Codes

- HTTP packets are sent along with status codes
- The status code is a standardized int value that tells you what happened
- This means that whatever language you are programming in, you can always expect a numeric code to tell you what happened

HTTP Status Codes

- 404 means file not found
- 403 means forbidden
- 401 means unauthorized (authentication required)
- 200 means success
- 418 means "*I'm a teapot*"
- More Status Codes
  http://en.wikipedia.org/wiki/List_of_HTTP_status_codes

ErrorDocument

- The ErrorDocument directive tells Apache how to handle certain errors.
- The ErrorDocument directive takes two parameters, an error code and a response (a message or a file)
- ErrorDocument allows us to specify an error and send back a custom html file.

ErrorDocument

- ErrorDocument reads file paths starting at your RIT www folder, so your paths start with a slash and your userid followed by the path from www.

  Example:

  ErrorDocument 404 /abc1234/error/error.html

  *your path must start with a / for this to work

Auth Workflow

1. Client requests protected resource
2. Apache sends:
    1. 401 Authentication Required header
    2. Realm: authentication name which is associated with the protected area of the web site.
3. Client Alert box pops up, displaying The Realm, where you type in the appropriate username and password for that realm.
4. Server verifies that the username in the approved list, and the password is correct. If so, send the resource to client.

Auth Types

- There are several types of authentication.
- The most basic one uses a password file, but it is not the most secure.

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess auth

AuthType shibboleth

AuthName "RIT"

ShibRequireSession On

SSLRequireSSL

require valid-user

.htaccess Auth

- Once logged in, how do you log out?
  - Clear browser cache

- Can I change the looks of the auth window?
  - No, but that's where web servers and logins come in

    Remember this is basic auth.

R·I·T

Exercise