How Web Servers Work

Lachele Foley

CHUGALUG 2014-05-27

Introductory Topic

Linux is widely used on the internet

- Roughly 1/3 run Linux
- Another 1/3 use a Unix OS
- As an aside
  - 96% of the TOP500 run Linux
    - TOP500: 500 fastest supercomputers
  - 46% of smartphones & tablets
  - 29% of embedded systems
- It's not just that it's free:
  - Open source = unlimited capabilities
- This is all I will say about the **Linux** part of LAMP

Stats: http://en.wikipedia.org/wiki/Usage_share_of_operating_systems

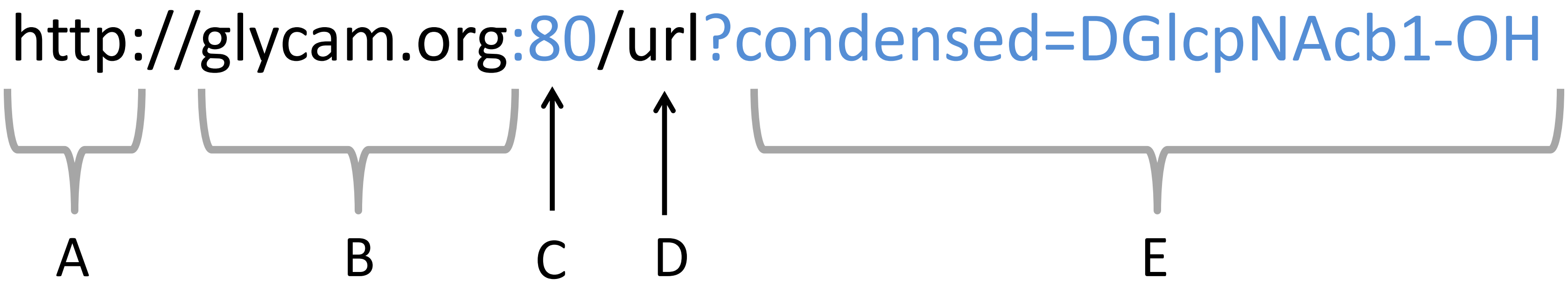
Ports

- Virtual network addresses inside your computer
- Like telephone menus, sort of
  - 80 is the standard "http" address
  - Here are some others:
    - 20-21 ftp
    - 22 ssh
    - 25 mail
    - 70 gopher (anyone remember?)
    - 80 http
    - 443 https
    - See /etc/services for more

- Apache (LAMP) listens for requests on port 80
  - There are other programs, but Apache is very popular

http://en.wikipedia.org/wiki/Port_(computer_networking)

Anatomy of a URL



http://glycam.org:80/url?condensed=DGlcpNAcb1-OH

- A: *Scheme* or *protocol*: service being requested
- B: *Host*: identifier of the website’s server
- C: *Port*: optional, usually omitted for websites
- D: *Path to resource*: location of info you want
- E: *Query String*: this is not part of the URL
  - It is passed on to some process on the server

When you click / hit enter

- Your URL becomes a request for information
- The http server returns information to you
  - The requested info if possible
  - An error message otherwise
- Your browser interprets and displays the information
- The information/content can be (among others):
  - Simple text
  - Text with markup (HTML)
  - A script to be run locally (JavaScript)
- Most of this happens *client-side*
- Let’s see some examples

Dynamically generated content

- Sometimes, the content can’t pre-exist
  - Search results
  - Shopping Carts
  - Facebook
- Software on the server (*server-side*) generates it
  - Pre-made, e.g., Wordpress
  - Written from scratch
- **Perl**, **PHP** and **Python** (LAM**P**)
  - These are popular languages
  - There are others
- See an example

Storing content for dynamic pages

- Consider all the blog posts, shopping carts, chats, events, greeting cards, search results, etc.
- Store in files on the hard drive?
  - If the info is relatively simple, maybe
- Better: use a database
  - Software queries DB for specific page info:
    - Background image, title, color scheme, main text
- **MySQL** and **MariaDB** (LAMP)
  - Again, there are others.
  - These two are popular (and closely related)