

# Chapter 22

## Web Hosting

# Web Hosting

## – Basics (1)

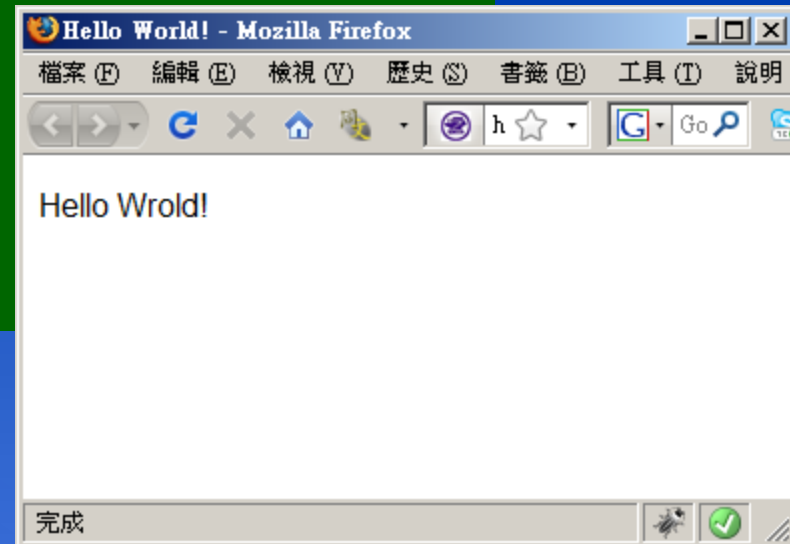
- Three major techniques in WWW (World Wide Web) System
  - > HTML
  - > HTTP
  - > URL
- HTML (1) – **H**yper**T**ext **M**arkup **L**anguage
  - > Providing a means to describe the structure of text-based information in a document.
  - > The original HTML is created by Tim Berners-Lee.
  - > Published in 1993 by the IETF as a formal "application" of SGML (with an SGML Document Type Definition defining the grammar).
  - > The HTML specifications have been maintained by the World Wide Web Consortium (W3C).
    - <http://www.w3.org/>

# Web Hosting – Basics (2)

## ◎ HTML (2)

- Mark-up the text and define presentation effect by HTML Tags.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">
<html>
  <head>
    <title>Hello World!</title>
  </head>
  <body>
    <p>Hello Wrold!</p>
  </body>
</html>
```



# Web Hosting

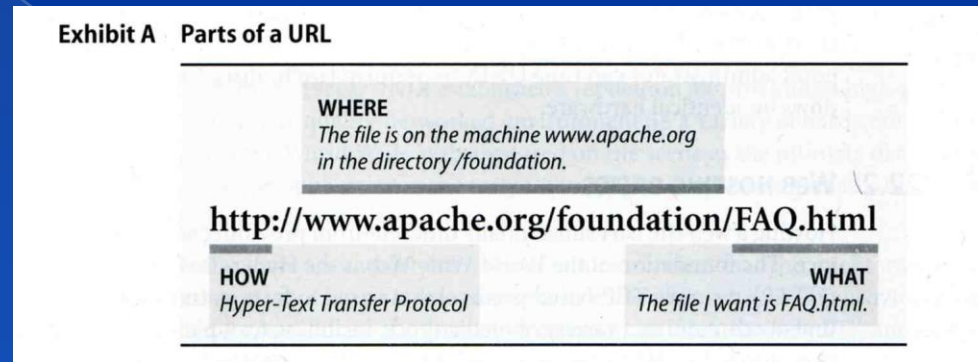
## – Basics (3)

- ◎ HTTP – Hyper Text Transfer Protocol
  - > A TCP-based protocol
  - > Communication method between client and server, both browsers and web servers have to follow this standard.
  - > Originally designed to transmit HTML pages.
  - > Now it is used to format, transmit, and link documents of variety media types
    - Text, picture, sound, animation, video, ...
  - > HTTPS – secured version.

[http://www.w3.org/Protocols/  
RFC 2616](http://www.w3.org/Protocols/RFC 2616)

# Web Hosting – Basics (4)

- ◎ URL – Uniform Resource Locator **RFC 1738**
  - > Describe how to access an object shared on the Internet



- > Format
  - Protocol:// [ [ username [ :password ] @ ] hostname [ :port ] ] [ /directory ] [ /filename ]
- > ex:
  - `http://www.cs.nctu.edu.tw/`
  - `ftp://ftp.cs.nctu.edu.tw/`
  - `telnet://bs2.to/`

# Web Hosting

## - Basics (5)

### ◎ URL Protocols

Proto	What it does	Example
http	Accesses a remote file via HTTP	<a href="http://www.cs.nctu.edu.tw">http://www.cs.nctu.edu.tw</a>
https	Accesses a remote file via HTTP/SSL	<a href="https://www.cs.nctu.edu.tw">https://www.cs.nctu.edu.tw</a>
ftp	Accesses a remote file via FTP	<a href="ftp://ftp.cs.nctu.edu.tw/">ftp://ftp.cs.nctu.edu.tw/</a>
mailto	Sends mail	<a href="mailto:lwhsu@cs.nctu.edu.tw">mailto:lwhsu@cs.nctu.edu.tw</a>
news	Accesses Usenet newsgroups	<a href="news:tw.bbs.comp.hardware">news:tw.bbs.comp.hardware</a>
telnet	Logs in to a remote computer	<a href="telnet://bbs.cs.nctu.edu.tw">telnet://bbs.cs.nctu.edu.tw</a>
file	Access a local file	<a href="file:///home/lwhsu/.tcshrc">file:///home/lwhsu/.tcshrc</a>

# Web Hosting

## - Client-Server Architecture (1)

### Client-server architecture

- > Web Server: Answer HTTP request
- > Web Client: Request certain page using URL

1. 以 URL 描述索取的資源位置向 Server 發送要求

3. 從 URL 描述的位置將 HTML 文件取出並回覆給 Client



5. 接收到 HTML 後由 Browser 解析後根據 HTML 描述定義將資料呈現出來

# Web Hosting

## - Client-Server Architecture (2)

- Using "telnet" to retrieve data from web server

```
knight:~ -lwhsu- telnet www.cs.nctu.edu.tw 80
Trying 140.113.235.111...
Connected to www.cs.nctu.edu.tw.
Escape character is '^]'.
GET /~lwhsu/hello.html HTTP/1.0
Host: www.cs.nctu.edu.tw

HTTP/1.0 200 OK
Date: Mon, 24 Nov 2008 21:40:29 GMT
Server: Apache
Last-Modified: Mon, 24 Nov 2008 21:30:27 GMT
ETag: "645b049-97-45c761a7ae2c0"
Accept-Ranges: bytes
Content-Length: 151
Content-Type: text/html
X-Cache: MISS from www.cs.nctu.edu.tw
X-Cache-Lookup: MISS from www.cs.nctu.edu.tw:80
Via: 1.0 www.cs.nctu.edu.tw:80 (squid/2.6.STABLE18)
Connection: close

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">
<html>
  <head>
    <title>Hello World!</title>
  </head>
  <body>
    <p>Hello Wrold!</p>
  </body>
</html>
Connection closed by foreign host.
```



# Web Hosting

## – The HTTP Protocol (1)

### ◎ HTTP: Hypertext Transfer Protocol

#### > RFCs: (HTTP 1.1)

<http://www.faqs.org/rfcs/rfc2068.html>

<http://www.faqs.org/rfcs/rfc2616.html> (Updated Version)

#### > Useful Reference: <http://jmarshall.com/easy/http/>

#### > A network protocol used to deliver virtually all files and other data on the World Wide Web.

- HTML files, image files, query results, or anything else.

#### > Client-Server Architecture

- A browser is an HTTP client because it sends requests to an HTTP server (Web server), which then sends responses back to the client.

# Web Hosting

## – The HTTP Protocol (2)

### ○ Clients:

※ Send Requests to Servers

#### > Action “path or URL” Protocol

- Actions: GET, POST, HEAD
- Ex. GET /index.php HTTP/1.1

#### > Headers

- Header\_Name: value
- Ex. From: someuser@jmarshall.com

#### > (blank line)

#### > Data ...

### ○ Servers:

※ Respond to the clients

#### > Status:

- 200: OK
- 404: Not Found
- ...
- Ex. HTTP/1.1 200 OK

#### > Headers

- Same as clients
- Ex. Content-Type: text/html

#### > (blank line)

#### > Data...

# Web Hosting

## - The HTTP Protocol (3)

action

Headers

status

Headers

Data

```
knight:~ -lwhsu- telnet www.cs.nctu.edu.tw 80
Trying 140.113.235.111...
Connected to www.cs.nctu.edu.tw.
Escape character is '^]'.
GET /~lwhsu/hello.html HTTP/1.0
Host: www.cs.nctu.edu.tw

HTTP/1.0 200 OK
Date: Mon, 24 Nov 2008 21:40:29 GMT
Server: Apache
Last-Modified: Mon, 24 Nov 2008 21:30:27 GMT
ETag: "645b049-97-45c761a7ae2c0"
Accept-Ranges: bytes
Content-Length: 151
Content-Type: text/html
X-Cache: MISS from www.cs.nctu.edu.tw
X-Cache-Lookup: MISS from www.cs.nctu.edu.tw:80
Via: 1.0 www.cs.nctu.edu.tw:80 (squid/2.6.STABLE18)
Connection: close

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">
<html>
...
</html>
Connection closed by foreign host.
```

# Web Hosting

## – The HTTP Protocol (4)

### ◎ Get vs. Post (client side)

#### > Get:

- Parameters in URL

GET <http://sabsd.cs.nctu.edu.tw/get.php?a=1&b=3> HTTP/1.1

- *No data content*

- Corresponding in HTML files

- Link URL: <http://sabsd.cs.nctu.edu.tw/get.php?a=1&b=3>

- Using Form:

```
<form method="GET" action="get.php"> ... </form>
```

#### > Post:

- Parameters in Data Content

POST <http://sabsd.cs.nctu.edu.tw/post.php> HTTP/1.1

- Corresponding in HTML files

- Using Form:

```
<form method="POST" action="post.php"> ... </form>
```

# Web Hosting

## – The HTTP Protocol (5)

### ◎ HTTP Headers:

#### > What HTTP Headers can do?

[Ref] <http://www.cs.tut.fi/~jkorpela/http.html>

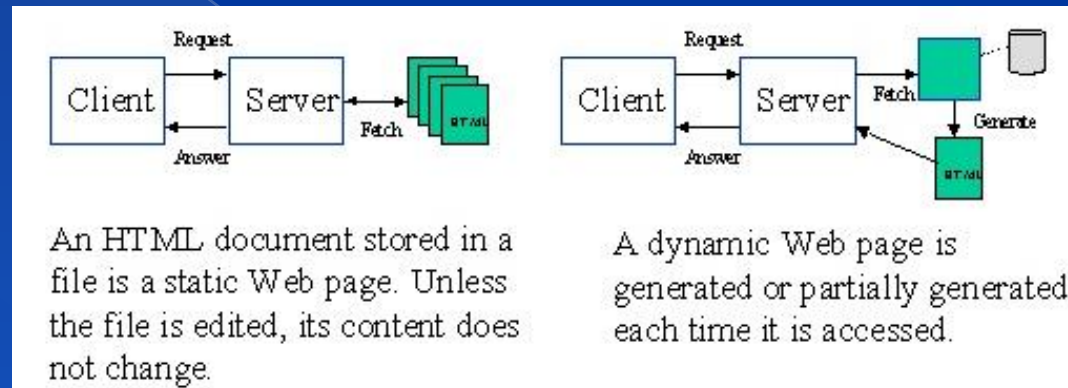
- Content information (type, date, size, encoding, ...)
- Cache control
- Authentication
- URL Redirection
- Transmitting cookies
- Knowing where client come from
- Knowing what software client use
- ...

# Web Hosting

## – Static vs. Dynamic Pages (1)



### Static vs. Dynamic Pages



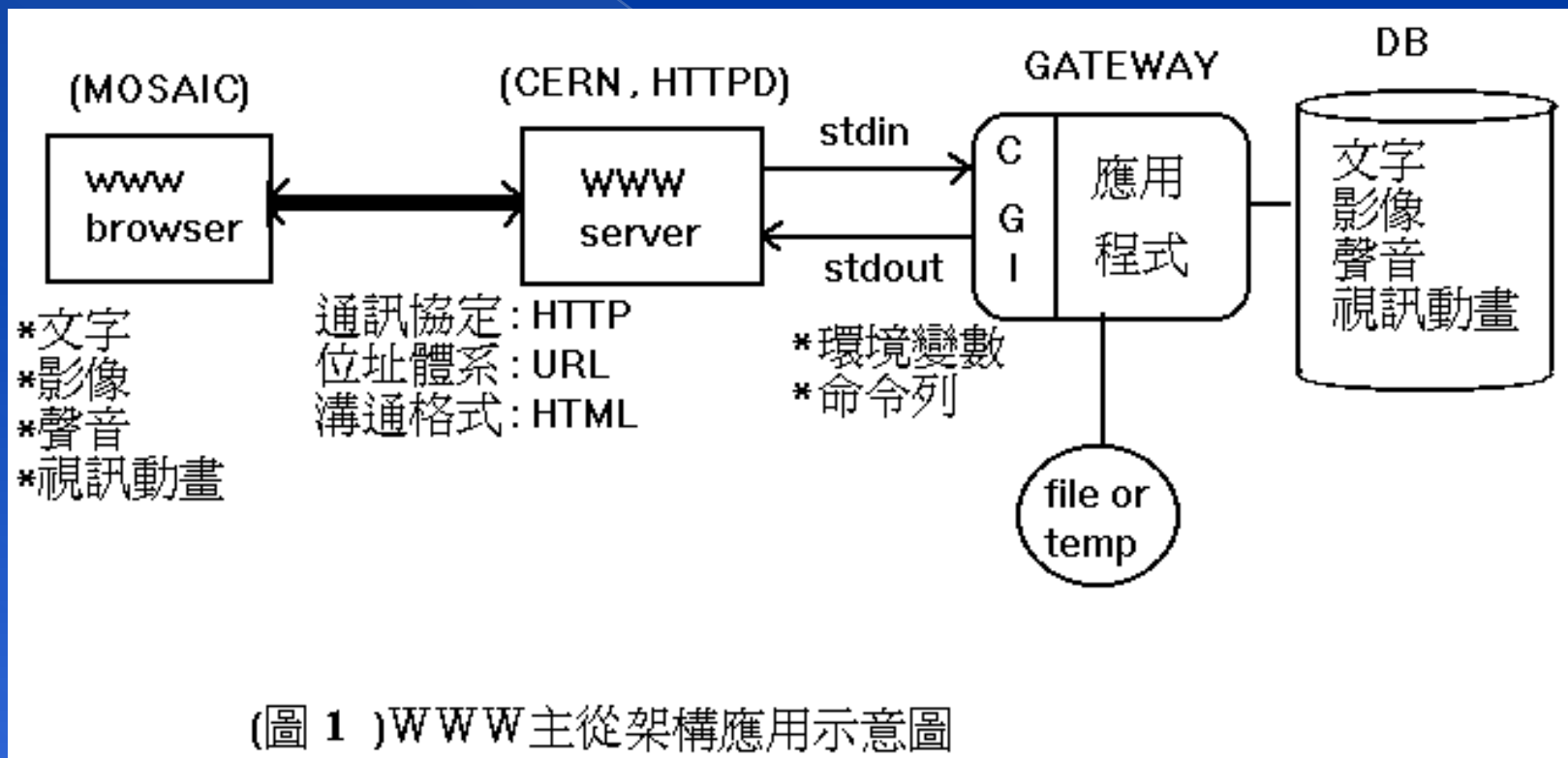
#### > Technologies of Dynamic Web Pages

- Client Script Language
  - JavaScript, Jscript, VBScript
- Client Interactive Technology
  - Java Applet, Flash, XMLHTTP, AJAX
- Server Side
  - CGI
  - Languages: Perl, ASP, JSP, PHP, C/C++, ...etc.

# Web Hosting

## - Static vs. Dynamic Pages (2)

- CGI (Common Gateway Interface)
  - > A specification that allows an HTTP server to exchange information with other programs



# Web Hosting

## – Virtual Hosts (1)

- Providing services for more than one domain-name (or IP) in one web server.
- IP-Based Virtual Hosting vs. Name-Based Virtual Hosting
  - IP-Base – Several IPs (or ports)
  - Name-Base – Single IP, several hostnames
- Example (Apache configuration)

```
NameVirtualHost 140.113.235.120

<VirtualHost 140.113.235.120>
  ServerName sysadm.cs.nctu.edu.tw
  DocumentRoot "/home/lwhsu/sysadm"
</VirtualHost>

<VirtualHost 140.113.235.120>
  ServerName netadm.cs.nctu.edu.tw
  DocumentRoot "/home/lwhsu/netadm"
</VirtualHost>
```

```
<VirtualHost 140.113.235.120:80>
  DocumentRoot "/home/lwhsu/netadm"
  ServerName sysadm.nctu.edu.tw
</VirtualHost>

<VirtualHost 140.113.17.221:80>
  DocumentRoot /www/tphp
  ServerName tphp.cs.nctu.edu.tw
</VirtualHost>
```



# Web Hosting

## – Virtual Hosts (2)

Q: How Name-Based Virtual Hosting works?

A: It takes advantage of HTTP Headers.

```
telnet cswproxy.cs.nctu.edu.tw 80
Trying 140.113.235.111...
Connected to www.cs.nctu.edu.tw.
Escape character is '^]'.
GET / HTTP/1.0
Host: www.cs.nctu.edu.tw

HTTP/1.0 301 Moved Permanently
Date: Tue, 25 Nov 2008 04:35:37 GMT
Server: Apache
Location: cht/announcements/index.php
Content-Length: 0
Content-Type: text/html

Connection closed by foreign host.
```

```
knight:~ -lwshsu- telnet
cswproxy.cs.nctu.edu.tw 80
Trying 140.113.235.111...
Connected to cswproxy.cs.nctu.edu.tw.
Escape character is '^]'.
GET / HTTP/1.0
Host: www.csie.nctu.edu.tw

HTTP/1.0 200 OK
Date: Tue, 25 Nov 2008 04:38:06 GMT
Server: Apache
Last-Modified: Thu, 15 Nov 2007 22:51:20 GMT
Content-Length: 113
Content-Type: text/html

<html>
<head>
<title>NCTU -- CS</title>
<meta http-equiv="refresh" content="0;
URL=http://www.cs.nctu.edu.tw/">

Connection closed by foreign host.
```

# Appendix: Virtual Interface

# Virtual Interface

## ◉ Idea

- > Let single machine responds to more IP addresses than it has physical network interfaces
- > Each of the virtual network interface can be associated with different domain name

```
FreeBSD: ~ -lwshsu- ifconfig
em0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=19b<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,VLAN_HWCSUM,TSO4>
ether 00:e0:81:4c:98:32
inet 140.113.17.209 netmask 0xffffffff broadcast 140.113.17.255
inet 140.113.17.207 netmask 0xffffffff broadcast 140.113.17.207
inet 140.113.17.208 netmask 0xffffffff broadcast 140.113.17.208
media: Ethernet autoselect (1000baseTX <full-duplex>)
status: active
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> metric 0 mtu 16384
inet6 fe80::1%lo0 prefixlen 64 scopeid 0x3
inet6 ::1 prefixlen 128
inet 127.0.0.1 netmask 0xff000000
```

# Virtual Interface Configuration – FreeBSD

## ◉ Using “ifconfig” command and “alias” option

- > `$ ifconfig em0 inet 140.113.17.210 netmask 255.255.255.255 alias`
- > `$ ifconfig em0 inet 140.113.17.210 delete`
- > `$ ifconfig em0 inet 140.113.17.210 -alias`  
(when you want to delete it)

## ◉ Specify in /etc/rc.conf

```
hostname="FreeBSD.cs.nctu.edu.tw"
```

```
defaultrouter="140.113.17.254"
```

```
ifconfig_em0="inet 140.113.17.209 netmask 255.255.255.0"
```

```
ifconfig_em0_alias0="inet 140.113.17.207/32"
```

```
ifconfig_em0_alias1="inet 140.113.17.208/32"
```