

Cơ bản về CGI

**Sở Thông Tin Truyền Thông Vĩnh Long
12/2012**

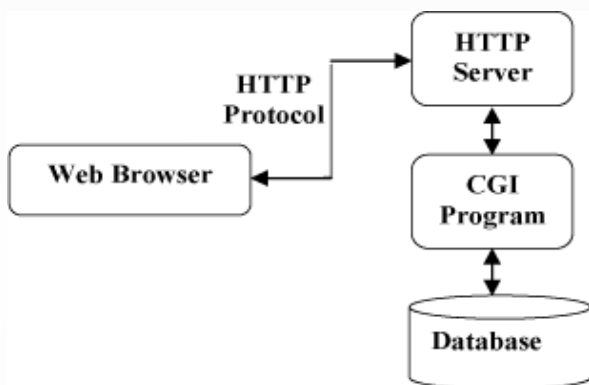
Trình bày: TS. NGÔ BÁ HÙNG
<http://ngobahung.vn>

Cơ bản về CGI

Nội dung

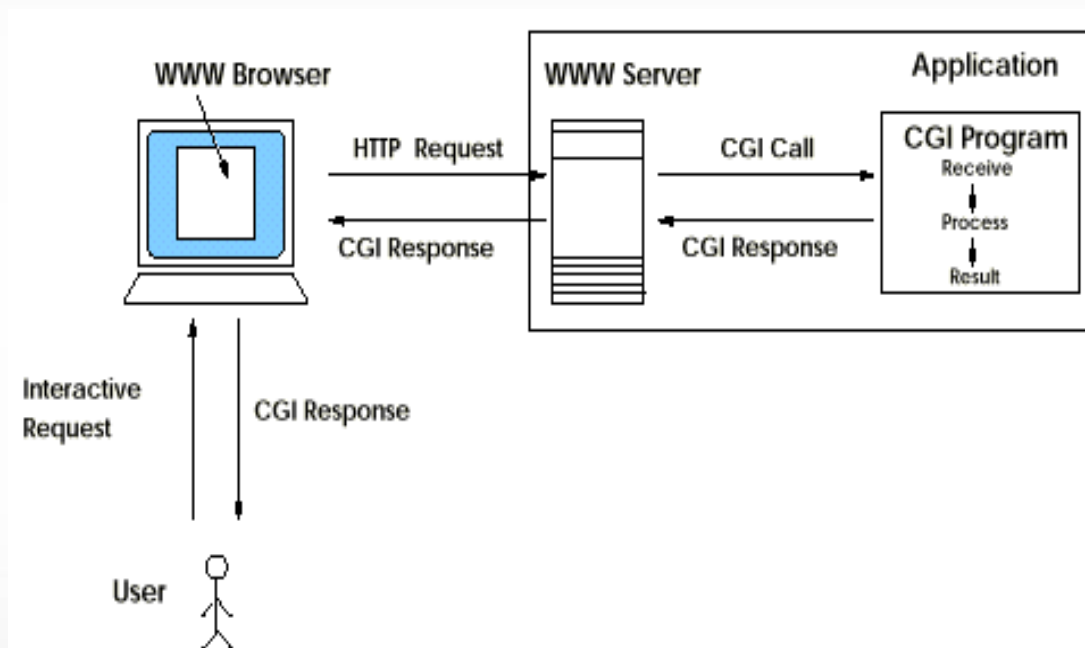
- CGI-Common Gateway Interface
- Chương trình CGI
- Các bước phát triển một ứng dụng nền Web với CGI
- Thực hành phát triển ứng dụng helloworld.cgi
- Các bước viết một CGI
 - Theo phương thức GET
 - Theo phương thức POST

CGI-Common Gateway Interface



- Là một qui ước về cách thức web server triệu gọi thực thi một chương trình và cách thức để chuyển tham số cho chương trình này
- CGI Program/CGI Script: là tập tin kịch bản (shell script, perl, ...) hoặc tập tin thực thi được (biên dịch từ chương trình nguồn)

Ứng dụng nền Web với CGI



Chương trình CGI

- Được ánh xạ đến từ một URL
- Tham số đầu vào
 - Dữ liệu nhập từ các Form HTML
- Dữ liệu xuất: Là các trang web động
 - Nội dung trang web được tạo mới cho mỗi yêu cầu người dùng
 - Xuất ra output chuẩn của chương trình CGI và được web server chuyển tiếp về Web Browser
- Có thể truy vấn/cập nhật cơ sở dữ liệu

Phát triển ứng dụng Web với CGI

- Cài đặt web server
- Cấu hình thư mục CGI trên Web Server
- Viết tài liệu HTML (FORM) để nhận dữ liệu nhập
- Viết CGI program/scripts được triệu gọi từ FORM
- Biên dịch CGI program
- Publish các trang HTML lên web server
- Đặt CGI program/script lên thư mục CGI của server
- Truy cập URL của HTML (FORM) để kiểm thử

Cài đặt và cấu hình Web Server

- Cài đặt web server Apache2 /Ubuntu
 - `sudo apt-get install apache2`
- Cài đặt web server Apache2 /Windows
 - <http://httpd.apache.org/download.cgi>
- Cấu hình apache2/Ubuntu
 - `/etc/apache2/sites-available/default`
 - Thư mục Document Root: `/var/www`
 - Thư mục CGI mặc định: `/usr/lib/cgi-bin`
 - Tài liệu: <http://httpd.apache.org/docs/2.2/howto/cgi.html>

Hello World CGI/C

```
// Lưu vào tập tin helloworld.c
#include <stdio.h>

int main(void)
{
    printf("Content-Type:text/html;charset=iso-8859-1\n\n");
    printf("<TITLE>Hello World</TITLE>\n");
    printf("<H3>Hello World !</H3>\n");
    return 0;
}
```

Biên dịch

- Mở terminal
 - `gcc -o helloworld.cgi helloworld.c`
- Copy chương trình helloworld.cgi lên thư mục **/usr/lib/cgi-bin**
 - Phải có quyền admin (sudo) để copy

Kiểm thử chương trình

- Mở trình duyệt web
- Truy cập đến chương trình
 - `http://localhost/cgi-bin/helloworld.cgi`

Thực hành

- Cài đặt, cấu hình web server Apache/Linux
- Viết chương trình helloworld.c
- Biên dịch chương trình helloworld.c thành helloworld.cgi
- Publish helloworld.cgi lên web server
- Truy cập đến helloworld.cgi từ web browser

Các bước viết một chương trình CGI

- Nhận dữ liệu đầu vào từ người dùng
- Lưu dữ liệu đầu vào vào các biến trong chương trình
- Cài đặt chức năng xử lý dựa trên các biến lưu dữ liệu đầu vào để đạt được những mục đích nào đó
- Tạo kết quả xuất dưới dạng một tài liệu HTML và gửi về người dùng

Dữ liệu đầu vào

- Người dùng nhập vào các Form trong tài liệu HTML
- `<form action="/cgi-bin/mult.cgi">`

Số 1: `<input name="m" size="5">
`

Số 2: `<input name="n" size="5">
`

`<input value="Nhập" type="submit">`

`</form>`

Nhập hai số:

Số 1:

Số 2:

HTML Form

- `<FORM ACTION=URL METHOD=POST/GET>`
 - `<input type="input-type" name="a-name">`
- `</FORM>`
 - ACTION: mô tả URL của CGI sẽ đón nhận dữ liệu của form
 - Method: Phương thức POST hoặc GET, mặc định là GET
- Input
 - Nơi người dùng nhập dữ liệu đầu vào
 - Có nhiều kiểu input tương ứng nhiều kiểu dữ liệu

Các thẻ của Form	Mô tả
<FORM ACTION="/cgi-bin/prog.pl" METHOD="POST">	Form Start
<INPUT TYPE="text" NAME="name" VALUE="value" SIZE="size">	Text Field
<INPUT TYPE="password" NAME="name" VALUE="value" SIZE="size">	Password Field
<INPUT TYPE="hidden" NAME="name" VALUE="value">	Hidden Field
<INPUT TYPE="checkbox" NAME="name" VALUE="value">	Checkbox
<INPUT TYPE="radio" NAME="name" VALUE="value">	Radio Button
<SELECT NAME="name" SIZE=1> <OPTION SELECTED> One <OPTION>Two ... </SELECT>	Dropdown List
<INPUT TYPE="submit" VALUE="Message!" >	Submit Button
<INPUT TYPE="reset" VALUE="Message!">	Reset Button
</FORM>	Form Ends

Gửi/nhận dữ liệu nhập

- Dữ liệu từ form gửi lên CGI theo một trong 2 phương thức GET hoặc POST, được mô tả trong thuộc tính method của thẻ Form: <Form ... method="...">
- Phương thức GET
 - Dữ liệu nhập được gán trong biến môi trường QUERY_STRING
- Phương thức POST
 - Được gán trong phần content của HTTP Request
 - Nhận vào từ ngõ nhập chuẩn (stdin) của CGI Program

Phương thức GET

- Dữ liệu của form được mã hóa và nối vào sau URL của CGI, ngăn cách bởi dấu ?
 - `http://servername/cgi-bin/mult.cgi?name=Ngo&last=Hung`
- QUERY_STRING
 - Là phần phía sau dấu ? chứa các cặp tên và giá trị (name=value) của các input được nhập bởi người dùng từ form
 - Các cặp tên và giá trị ngăn cách nhau bởi dấu &
 - Ví dụ: `name=Ngo&last=Hung`

Phương thức GET (tt)

- Chương trình CGI sẽ phải
 - Đọc biến môi trường QUERY_STRING
 - Phải biết trước tên (name) các input của form
 - Phân tích giá trị QUERY_STRING để tách ra các cặp {tên, giá trị}
- Đối với chương trình CGI viết bằng C
 - Hàm `char* getenv("QUERY_STRING")` sẽ trả về giá trị của biến môi trường QUERY_STRING
 - Hàm `sscanf()` lấy giá trị các input từ QUERY_STRING

Ví dụ về GET method

```
// Chương trình mult.c
#include <stdio.h>
#include <stdlib.h>
int main(void)
{
    char *data;
    long m, n, result;
    data = getenv("QUERY_STRING");
    sscanf(data, "m=%ld&n=%ld", &m, &n);
    result = m*n;

    ....
}
```

Nhập hai số:

Số 1:

Số 2:

Tạo kết quả trả về

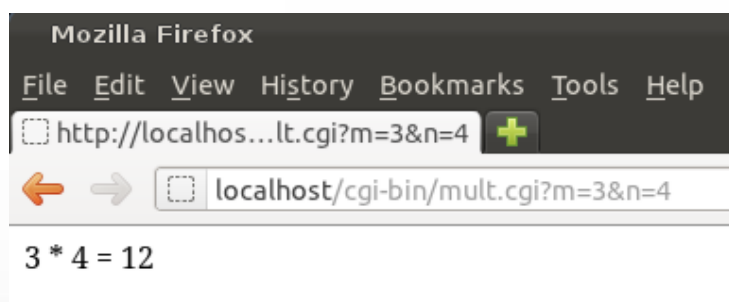
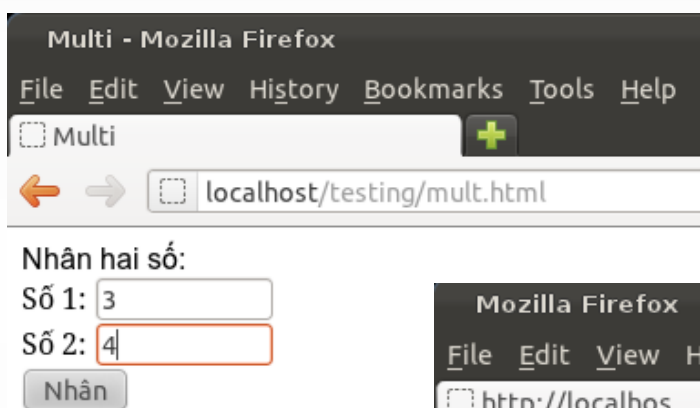
- In ra thiết bị xuất chuẩn theo cấu trúc sau

status line (protocol status code status phrase)	→	HTTP/1.0 200 OK
header lines	→	Server: Apache/1.3.0 (Unix) Content-Length: 6821 Content-Type: text/html
data, e.g., requested html file	→	<HTML> <BODY> ... </BODY> </HTML>

Tạo kết quả trả về (tt)

```
// Tiếp tục chương trình mult.c
data = getenv("QUERY_STRING");
sscanf(data, "m=%ld&n=%ld", &m, &n);
result = m*n;
printf("Content-Type:text/html\n");
printf("\n");
printf("<HTML><BODY>\n");
printf("%ld * %ld = %ld", m, n, result);
printf("</BODY></HTML>");
}
```

Kết quả thực hiện



Phương thức Post

- Đặt method=POST trong form
- Dữ liệu nhập từ form sẽ được chuyển lên server web sau khi tất cả các header đã gửi đi
- Dữ liệu nhập từ form được chuyển cho chương trình CGI thông qua kênh nhập chuẩn (stdin) của chương trình CGI
- Kích thước của dữ liệu nhập (tính bằng bytes) cũng được gửi cho chương trình CGI thông qua biến môi trường CONTENT_LENGTH

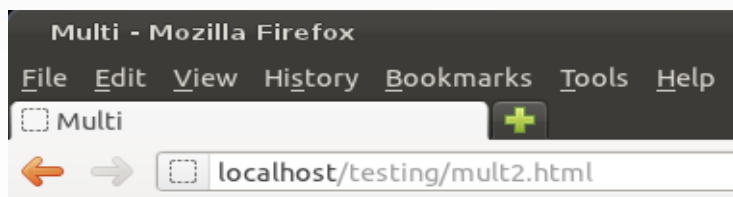
Phương thức Post (tt)

- Chương trình CGI viết bằng C phải
 - Lấy chuỗi CONTENT_LENGTH=xxx chứa chiều dài của dữ liệu nhập từ form từ biến môi trường, dùng hàm `getenv("CONTENT_LENGTH")`
 - tách số bytes xxx từ chuỗi CONTENT_LENGTH=xxx, dùng `sscanf(strlen, "%d", &len)`
 - Đọc dữ liệu nhập từ kênh nhập chuẩn stdin, dùng hàm `fgets(data, len+1, stdin);`
 - Tách giá trị của các input từ form gửi lên, dùng hàm `sscanf(data, "m=%ld&n=%ld", &m, &n);`

Ví dụ về phương thức Post

```
// Chương trình mult2.c
#include <stdio.h>
#include <stdlib.h>
int main(void)
{
    char *strlen; char data[200];
    long m,n,result; int len;
    strlen = getenv("CONTENT_LENGTH"); // Lay chieu dai du lieu nhap
    sscanf(strlen,"%d",&len);           // Doi chuoai thanh so
    fgets(data, len+1, stdin);          // Doc du lieu tu thiet bi nhap chuan
    sscanf(data,"m=%ld&n=%ld",&m,&n);
    result = m*n;
    ... // Phần kết quả trả về giống như ví dụ về phương thức GET
```

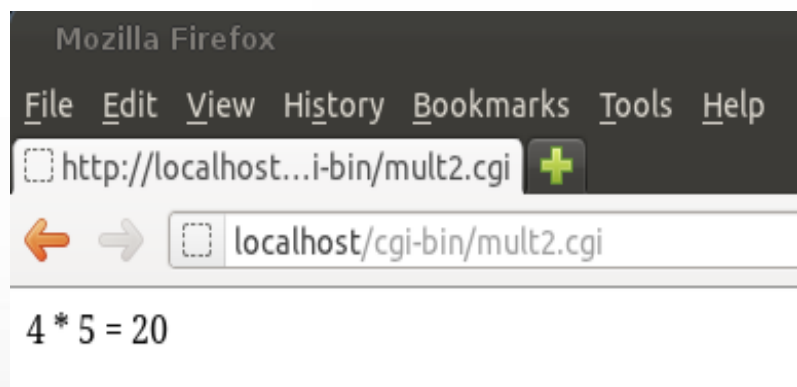
Ứng dụng với phương thức POST



Nhập hai số:

Số 1:

Số 2:



Bài tập thực hành

- Tạo tập tin mult1.html chứa form như hình bên với phương thức là GET và gọi CGI phía dưới đây
- Viết CGI bằng C có tên là mult1.c
 - Đón nhận dữ liệu từ mult1.html
 - Thực hiện nhân 2 số và trả kết quả về cho người dùng
- Publish và kiểm thử CGI trên từ trang web mult1.html

Nhân hai số:

Số 1:

Số 2:

Bài tập thực hành

- Tạo tập tin mult2.html chứa form như hình bên với phương thức là POST và gọi CGI phía dưới đây
- Viết CGI bằng C có tên là mult1.c
 - Đón nhận dữ liệu từ mult2.html
 - Thực hiện nhân 2 số và trả kết quả về cho người dùng
- Publish và kiểm thử CGI trên từ trang web mult2.html

Nhân hai số:

Số 1:

Số 2:

Biến môi trường

- Được Web server thiết lập khi gọi thực thi một CGI
- Được CGI tham khảo đến để tạo ra các output
- Có hai loại
 - Chung cho tất cả yêu cầu: Các biến này được thiết lập cho mọi yêu cầu
 - Riêng cho từng yêu cầu: Được thiết lập tùy thuộc vào từng yêu cầu riêng biệt cho từng CGI riêng

Một số biến môi trường

SERVER_NAME

- Tên hoặc địa chỉ IP của server

SERVER_SOFTWARE

- Tên và version của phần mềm làm server web đang trả lời Client

SERVER_PROTOCOL

- Tên và version của giao thức mà Request đã được truyền đến

REQUEST_METHOD

- Phương thức mà Request đã được truyền tới

QUERY_STRING

- Dữ liệu nhập được nối vào URL sau dấu "?"

Một số biến môi trường (tt)

CONTENT_TYPE

- Chuẩn mã hóa dữ liệu , ví dụ "text/html"

CONTENT_LENGTH

- Chiều dài dữ liệu nhập gửi vào chương trình CGI thông qua kênh nhập chuẩn của CGI

HTTP_USER_AGENT

- Trình duyệt mà client đang dùng

DOCUMENT_ROOT

- Thư mục gốc của Web server

Ví dụ về biến môi trường

SERVER_SOFTWARE = Apache/1.3.14

SERVER_NAME = www.ncsi.iisc.ernet.in

GATEWAY_INTERFACE = CGI/1.1

SERVER_PROTOCOL = HTTP/1.0

SERVER_PORT = 80

REQUEST_METHOD = GET

HTTP_ACCEPT = 'image/gif, image/x-xbitmap, image/jpeg, */*'

SCRIPT_NAME = /cgi-bin/environment-example

QUERY_STRING =

REMOTE_HOST = ece.iisc.ernet.in

REMOTE_ADDR = 144.16.64.3

Câu hỏi & Thảo luận

Môi trường thực hành (1)

- Cho các máy Ubuntu
 - Theo hướng dẫn của Slides
- Đối với máy tính Windows
 - Download phần mềm putty.exe từ địa chỉ <http://172.16.66.39/Softs/putty.exe>
 - Truy cập đến server 172.16.66.39, có 2 thư mục chia sẻ
 - wba-html: lưu các trang web của bạn
 - Wba-cgi: Lưu các chương trình CGI của bạn

Môi trường thực hành (1)

- Soạn chương trình CGI bằng ngôn ngữ C với một trình soạn thảo văn bản bất kỳ (Notepad)
- Copy chương trình lên thư mục wba-cgi
- Chạy chương trình putty.exe
 - Nối kết đến server 172.16.66.39, port 22
 - User: chk17, password: chk17
 - Chuyển vào thư mục /usr/lib/cgi-bin/wba
 - Biên dịch gcc -o hello.cgi hello.c
- Copy tập tin HTML lên thư mục wba-html
- Địa chỉ của trang web: <http://172.16.66.39/wba/mult.html>
- Địa chỉ của cgi: <http://172.16.66.39/cgi-bin/wba/hello.cgi>