Server Side Story

Nigel Whitfield answers a query on Server Side Includes, enabling you to replace parts of your web page with other information. But first, he comments on plans to net-connect all UK schools.

So, with the help of Bill Gates, every school in the country is to be connected to the internet. And almost unanimously, this has been hailed as a "good thing". The internet is, after all, a good thing, so giving everyone access to it from school must be as well, mustn't it?

Perhaps. But there's also a danger that, rather than giving everyone equality of access, making the net available in every school will actually increase the divide between different classes of pupil. Rubbish, you might say; and if all it's used for is a research tool, then there's some merit in that comment.

After all, those who don't have a computer, or lack net access at home, can always visit a library to look things up, or rely on the old family encyclopaedia, can't they? But what happens when the net ceases to be just a tool and is an accepted part of education, as some would like to see it? When the object of the exercise becomes not "Find me ten useful facts about life in Uruguay" but "Find me ten web sites about Uruguay", all those whose access is restricted to their time at school will find themselves at a disadvantage compared to children from families with access at home, where they can search out new sites at their leisure.

There are some solutions, of course. After-school clubs will allow some children access to the computers at school that they didn't get to use during the day, and placing terminals in libraries will help as well.

What's really needed, however, if acceptance of the internet is to become anything more than another divide between the haves and the have-nots, is for access to be simple, for the "internet appliance" to become reality, cheaply and easily. It's no good to say that a computer is cheap, and that a net connection is only ten pounds a month when that's still far more money than many people can afford.

If the internet really is to become pervasive, in schools as well as homes, it needs to be cheaper and easier to access. Forget Windows 95, and the combined browser and desktop of IE4. What's needed is a leap as significant as the arrival of the first Mac into the DOS-based world of the eighties, at a price that's comparable to a video recorder and just about as simple to use.

Only when people can rent a net applicance as easily and cheaply as they can a television set, will the internet really be a tool that we can wholeheartedly welcome into schools. Of course, putting every school on the internet is a start, and it will doubtless prove to be a valuable resource for many people. But it's vitally important to keep a sense of proportion and remember that it's not an unqualified advantage to both teachers and pupils, nor necessarily the best use of funds.

Rather than rushing headlong towards the utopian view of the net shared by many users, perhaps it's better that those of us who use it already think about what we're doing and the effect it will have on all parts of society, lest we create a divided dystopia.

Platform tickets with Server Side Includes

Most web pages that people create tend to be static, or to rely on multimedia and fancy add-ins to make them look more attractive to people. While that's all very useful, what about including information that will depend on who's accessing your site, or other information that changes from time to time.

As one of the questions this month shows (page 250), it's fairly easy to add to your web site the time a page was changed, but the Server Side Include (SSI) mechanism (also known as server parsed HTML) can do more than that: you can create pages that might include a reference to where someone's connected from, or the type of computer they're using, or even present a different menu completely, for some types of user. A word of warning: While SSI pages can be very useful, they can also put an extra load on your server. As a result, it's common to tell the server only to look at pages with a particular extension, usually **shtml**.

At its simplest, you can use the technique to add a standard button bar to every page. Then, if you want to give your site a makeover, just change the file that contains the button bar, and every page that uses it will look different. Commands are embedded in HTML comments, so you'll either need to edit your web pages manually, or use a web editor that allows you to put in comments. For example, the code:

```
<!–#include virtual="/common/buttons.html"–>
```

would insert the code contained in **/common/buttons.html** where the comment appeared in your web page. If you replace "virtual" with "file" then you can specify the exact location of a file anywhere on the server, rather than within the main web directories.

You can also display any of the information passed to the web server by the browser, or maintained internally, using the **echo** command to display a variable. For instance:

```
<!–#echo var="REMOTE_HOST"–>
```

will be replaced with the name of the computer that's accessing your web

page, so put it in the middle of a 'Hello and welcome' message for a personalised touch. There are a few other commands too; include **flastmod** and **fsize** which tell you the date last changed and size of a file, respectively. But one of the most useful is the **exec** command which our example uses to create a simple page that will display different information for Mac users.

Here's the basic code for the page (missing the HEAD portion):

```
    <h1>Welcome to my test page</h1>
```

The "echo" command can help you personalise your site for the Mac (left)...

...and for Windows 95 PCs with a minimum of coding (right)

```
1   Your browser is <!–#echo var="HTTP_USER_AGENT"–><P>
    The next lines will change depending on which platfom you're
using<P>
2   <!–#exec cmd="/usr/local/etc/httpd/cgi-bin/platform_ticket"–><P>
        And now we're back to the rest of the file.
```

The first marked line simply displays the information returned by the user's browser, which is sent to the web server and stored at HTTP_USER_AGENT; the second line calls a program called "platform_ticket" which is a Perl script. Instead of saying "cmd=" you could say "virtual=" and specify the program as a name like "/cgi-bin/platform_ticket", but on some servers (we used Netscape FastTrack) you might not be able to access the files you need from within the script unless you give the full path to them. Don't forget you'll need to include code like this in a file ending in .shtml and enable Server Side Includes on your web server. That option may not be available if you simply have free space with your internet account.

The next thing to do is to create the different versions of the page that you want to display for Windows and Mac users, and save those sections in files called "win.inc" and "mac.inc" (or whatever you've referred to in your script). Finally, you'll need to write the actual script itself, which decides which file to send back to the browser:

```
        #!/usr/local/bin/perl
        # script to output a file based on the browser accessing a site
1       $browser = $ENV{'HTTP_USER_AGENT'} ;
2       if ( $browser =~ /mac/i ) {
3             open( INCLUDE, "mac.inc" ) ;
4       } else {
           open( INCLUDE, "win.inc" ) ;
        }
5       while( <INCLUDE> ) {
           print $_ ;
        }
6       close INCLUDE ;
```

1. First, we've saved the information into a new variable; not strictly necessary, but if you wanted to change the value, you could do so more safely. 2. This line checks to see if $browser contains the string "mac", and the i on the end ensures that it's case insensitive. 3. If "mac" occurs in the $browser variable, then the file mac.inc will be opened; INCLUDE is the "handle" given to the file so we can refer to it later. 4. To do more tests, change "else" to "elsif" and add more conditions here. 5. These next three lines read from the file that was opened, one line at a time, printing it out. 6. Finally, we close the file, and the job's done. So, with just a handful of lines of Perl, and a couple of comments in your web page, you can give people the feeling that they've visited a site that's a little more customised to their needs; and of course you can do more than just display pages based on their browser. For more information about Server Side Includes, visit twister.luton.ac.uk/Manual/ssi.html.

p250 ➢

Questions & Answers

Q We've created an intranet based on a system running Netscape's FastTrack server. So that everyone knows when a page has been updated, we'd like the server to include information automatically. I've heard of Server Side Includes, but can't find much information about them. How do we do this?

A Server Side Includes, also referred to in the Netscape documentation as parsed html, are a way of telling the server to replace certain parts of a page with other information; if you decide to use this, you'll have to tell the server which pages you want it to parse. The usual way is to save pages with the extension **.shtml** and include commands within them, embedded in HTML comments, which allow you to add features such as the date a file was modified or the name of the system requesting the page, and so forth.

The panel on page 248 gives details of how you can use Server Side Includes to create pages that will depend on the type of computer being used to access them. However, if all you want is a simple date and time on the bottom of your pages, Netscape FastTrack server has a "page footer" feature that you can use to achieve the same result, without having to worry about server side includes.

To activate custom page footers, use the server manager, and click on the Content Management button in the top frame, then choose Document footer in the left pane, and you'll see a screen similar to the one displayed here. You can type the footer text that you want in the box, including the tag **:LASTMOD** for the last modification date of the file, and select the date format from the drop down menu.

When you've made your changes, click on OK, then choose the button marked Save and Apply, stop, then restart your web server. All your pages, unless you selected just a portion of the server, will now have a footer added automatically.

Netscape's FastTrack server can automatically add a timestamp to all your pages

Q I am an AOL user and I want to know how I can connect to the IRC network. Do I need to configure my system specially?

A No, you don't need to do anything special if you have the latest version of the AOL software, although older versions won't work with programs that require a 32-bit winsock stack (the part of Windows that links you to the internet).

The best thing to do is to upgrade to the latest version of the AOL software (which you can do at keyword UPGRADE). If you have it already, the system will tell you when you try to download it again. When you've installed AOL, you'll find that in one of the AOL directories there's a file called **winsock.dll**; install your IRC program into the same directory (or alternatively, make sure that it's in the path) and then all you need to do to connect to IRC is to start AOL, then start your IRC program.

One of the best IRC programs to use is called mIRC; you can download it from www.mirc.co.uk or at keyword IRC on AOL, where you'll also find additional information about configuring your system.

Q For the purposes of web-page design, what safe assumptions can I make about resident fonts? Is there a standard set installed with Windows 3.1 and 95 and, if so, where can I learn what it is?

A First, don't forget Mac users! The best thing to do is to assume as little as possible about fonts, as different systems may have widely differing selections. The best bet is to stick to the core TrueType fonts, which includes Arial, Courier New and Times New Roman; Mac users who have installed Microsoft Office applications will almost certainly have these fonts available. Those who don't can download them (and others) from www.microsoft.com/truetype/fontpack/mac.htm.

If you really do want to use lots of fonts to create a specific look for your web site, the best bet may be to provide links to the Microsoft (or other) font pages from the front page of your site, so that those who want to will be able to download them and make sure that they see things as you really intend them.

Q I am using Internet Explorer 4 Final (Build 4.71.1712.6) and MS Dialup Networking 1.2 on Windows 95 4.00.950. When I want to browse the internet it won't dial automatically. I think the reason is because the Save Password checkbox is greyed out. I have tried reinstalling the connection several times following the help exactly, but to no avail. I have also tried using Netscape Communicator but, again, I still cannot check the checkbox.

A Windows doesn't think you're signed on to the system as a valid user. When you install networking, Windows 95 assumes that you need to log on before the system is secure. However, many people simply press the Escape key when they're asked for a password, and since everything else works properly, don't realise the importance of what they're doing.

When you start Windows, if you don't

Setting a blank password will stop Windows prompting for a user name

p252 ➢

have a user created, give yourself a name and then leave the password blank. You'll be asked to confirm it, so leave the box blank again. Now, when you start the system, you won't be asked for a user name.

If you've already created a password and want to avoid having to say who you are each time you start the system, you need to log in properly when you start up, with the correct password (or create a new user, with no password, instead), open the control panels and choose Passwords, then set the password to be blank. You should be able to tell Dialup Networking to save the password, so you'll be able to start your computer and get on the internet, without having to remember anything more than where the power switch is.

Q I'm writing CGI scripts in Perl 5. As Perl and most scripts/libraries are freely available on the net, what's the position regarding copyright?

A Your scripts are your own intellectual property, and unless you decide to give them away, then there's nothing anyone else can do with them. But if you were to distribute, for example, a Perl-based online shopping system which used freely available libraries from the internet, you may find that the conditions of use of those libraries prohibit you from using them as part of your commercial product. After all, why should you sell something someone else wrote?

It's most unlikely that someone would be able to steal your scripts, unless the security on your web server is more or less non-existent. If the scripts are kept in a script directory, it's hard for anyone to see them. All that's passed to the web browser is the results of running the script itself.

You should clearly mark your scripts with your name, and other information, anyway, including an indication of your copyright, and a clear prohibition on use without your permission, if you wish to restrict their use, though of course it can be very hard to prove that someone has used your script, since it can easily be modified to look very different (especially with a language as flexible as Perl).

Q I'd like to set up a dial-in PPP link to the Unix server in my office, so that I can connect to the internet that way, rather than paying to subscribe to an ISP. How do I allocate IP addresses to the link? One of the manuals says something about having to have a separate subnet; what's that, and how do I create one? Is it absolutely necessary?

A A subnet is a section of addresses in a range of TCP/IP network numbers that, when written in binary, has the top section of bits all the same. For instance, if you have a network of 256 addresses, you might use four bits to create subnets, giving you sixteen subnets, each capable of having fourteen machines on it (addresses with the bottom bits all either 1 or 0 aren't used).

By putting your PPP link on a subnet, the other computers on the network will be able to work out easily that traffic for that subnet has to go via a certain route. However, that's not always possible, and it might mean re-configuring lots of your systems just to add a single dial-in link.

An alternative solution is to use proxy-ARP. ARP is the Address Resolution Protocol, and it's the way in which computers running TCP/IP match ethernet addresses to IP numbers. With proxy ARP, you add an entry to the arp table (usually using the /etc/arp command) that effectively means your Unix server is saying to the rest of the network "If you've got anything that's destined for this address, send it to my ethernet card." When the information is received at the Unix server, it'll be passed to the PPP interface.

Since you don't say which version of Unix or PPP you're using, it's hard to be more specific, but it looks like using proxy ARP may be the best solution to your problem. Most modern versions of PPP will be able to make the arp entry automatically when they receive a connection; check your manual.

PCW Contact

Nigel Whitfield is a freelance journalist, maintainer of several internet mailing lists and consultant to a number of non-profit organisations. He can be contacted via internet@pcw.co.uk.

Command **post**

Nigel Whitfield has advice and a useful list of basic Unix commands, to help you manage your web-space files, and tells you which tools you will need to make the best use of them.

It's not too long ago that real audio was the latest advance in multimedia over the internet. Then came a few video systems (including real video and VDO) generally enabling corporate or campaigning videos to be viewed, without the delay of first waiting for files to be downloaded.

There have been concerts on the internet, such as the Rolling Stones on the MBone, but the idea of simultaneous broadcasts for news events is pretty new — and unthinkable until only a year or so ago. Yet events surrounding the death of Diana, Princess of Wales, were available with live TV coverage on the internet, or radio commentary for those who didn't have the bandwidth or the software necessary to view pictures. Many people first learned of the tragedy from the splash screens, as they logged on to services such as AOL in the morning to check their email.

Just as there have been other events, like the 1952 Coronation, that have seen a particular medium show itself in a new light to many people, so the internet in this instance may have appeared to be a medium capable of providing up-to-date news and information on demand. It was not a "defining moment", though. While some people will have viewed the coverage, and others been amazed at the breadth of information available so quickly, there is little doubt that, for many the limitations of the current technology will have been only too apparent, as connections were dropped or timed out. The net may have proved itself a strong competitor to other media, but it has also shown the extent to which it desperately requires more capacity if it is to deliver the right message.

Capacity is one of those things that everyone is always wanting more of; and while some of the major backbone providers are making massive upgrades, for many of us, the first step is going to be to upgrade our modems to a faster standard.

While 33,600bps is now established, it's still not ideal for many people. If you're doing lots of surfing, looking at media-rich pages, you'll probably be considering upgrading to one of the new 56K modems; and, as you will read in one of this month's questions (*pp244/245*) people are already disappointed with the performance.

The simple truth is that the 56Kbps figure will be seen by some people only in ideal conditions. And for those who work from home, for whom the ability to send information fast is equally as important, 56K offers no benefits.

The next stop is ISDN which has been heavily promoted in the press by BT, with a discount offer to save you money on the installation. It's a shame that BT doesn't work in quite the same way with call charges. Heavy modem-users are already likely to use PremierLine to receive discounts on their calls and, if you have two lines at home, one fee will apply the discounts to both lines — unless, of course, one of them is ISDN.

BT still seems incapable of accepting that people might have an ISDN line at home: you can't order it as a residential service, and because of that, PremierLine isn't available. You could choose the similar business discount scheme, but you'll have to pay a separate membership fee for that. So while a business user can have all their lines discounted for a single fee, home users are still discriminated against. Given that British Telecom's internet organisation currently offers ISDN access for the same price, wouldn't it be nice if the phone company could do the same, or at least refrain from such blinkered discrimination against home users who want to make the most of the net?

Managing files on a Unix server

Despite the popularity of Windows NT, chances are that if you have web space on a server run by an internet provider, it's provided using the Unix operating system. `

While many providers allow you just to upload your files via ftp, others actually allow you access to the server, so you can rename files or perform other housekeeping tasks on your web space more easily. To do that, though, you'll need a grasp of some of the basic commands that a Unix system provides. And once you have the hang of those, you'll be able to make better use of the space and save yourself a lot of time: for instance, by uploading a Unix-style TAR (Tape ARchive) archive of your site as a single file and unpacking it at the other end.

So, what are the basic commands, and what tools will you need to make use of them? First, Telnet is the program that you need to connect to the web server: your provider will tell you what machine to connect to, and if you're running Windows 95, there's a Telnet program included. Mac users can download a copy of NCSA Telnet from most well known archive sites and various versions of the program are readily available for other platforms.

The other useful tool is the TAR program; TAR is a format for collecting a whole set of files into one large file, rather like a Zip archive but without compressing the information. Some providers do have versions of Zip on their Unix systems, but TAR is universal. MacTar, PCTar and other versions for different platforms can be obtained from your favourite software library.

Once you've connected to the web server via Telnet and logged in with the password supplied, you'll be presented with a prompt. It might be a $ symbol, a %, or just about anything else; you can now type commands to the Unix operating system in much the same way that you'd control a PC running MSDOS.

The first thing to remember is that, just as in a URL, the slashes that separate directory names in Unix are forward (/) rather than the backwards slash (\) used on a PC. So if you have a sub-directory in your web space called "documents" with another one named "sales", you can refer to the latter from your main directory by referring to **documents/sales**.

Just like on a PC, you can change to a directory using the **cd** command, followed by the name of the directory: **cd documents/sales**, for instance. However, unlike DOS, the cd command on its own doesn't tell you what directory you're in. Instead, you need to type **pwd** (Print Working Directory) because cd on its own will return you to your home directory, the one that you're placed in when you first connect.

If you want to rename a file, for instance, to keep your old index file as **old_index.html**, in Unix parlance you move it. You can even rename it into a different directory. Just type **mv index.html old_index.html**, or if you have a directory called old, try **mv index.html old** and it will become **old/index.html**.

And how to make the "old" directory? All you need to do is type **mkdir old**. To remove it again, type **rmdir old** (as long as you've removed all the files it contained). You can remove just one file with the remove command, **rm**. You can also remove a directory and the files in it, just like the DOS **deltree** command, by adding the "recurse" switch to the command. In our example, the "old" directory and all its contents can be deleted by typing **rm -r old**.

To copy a file, the command you need is **cp**; just as in DOS, you can specify a directory as the second option, but if you want to copy a file into the current directory, you must specify it as '.' rather than missing it off, like **cp …/incoming/index.html**.

Finally, to unwrap an archive file created using a version of TAR, such as **myweb.tar**, into the current directory, you need to type **tar xvf myweb.tar**. The x stands for extract, the v for verbose, so you can see what's happening, and the f means that you'll be giving the name of the archive file. There are a few more commands in the table (*right*), and these should be enough to do most of the tasks you'll need to perform in order to manage your web space. ■

Handy Unix commands for managing your web space

Type:	…in order to:
cat file	view the contents of file
cd	change to your home directory
cd directory	change to directory
chmod 600 file	make file accessible only by yourself; this may prevent the web server from reading it, depending on how it's configured
chmod 644 file	make file accessible by everyone
cp file1 file2	copy file1 to file2
cp file1 directory	copy file1 into directory
exit	finish your current session
ls	list the contents of the current directory (you may specify another directory, or a file)
ls -l	detailed listing of the current directory
more file	view the contents of file a screen at a time; press space for the next screen
mv file1 file2	rename file1 as file2; file2 can include a directory name
mv file1 directory	move file1 into directory
pwd	display the name of the current directory
rm file	delete file
rm -r directory	remove directory and all its contents
rmdir directory	remove directory (which must be empty)
tar xvf file	unpack the archive file into the current directory
tar cvf file .	make an archive called file of the current directory and all its sub-directories

Wildcards can be used in file names, especially where the second option is a directory. * represents more than one letter, and ? a single letter. For example:

mv *.html old	moves all html files into the "old" directory
cp ?.gif oldgifs	copy a.gif, 1.gif and so on, but not 23.gif to the oldgifs directory
.	is shorthand for the current directory
..	is the next directory up
~	is your home (login) directory

p244 ➢

Net.questions and answers

Q I have registered a domain name with a UK company which offers that service, and I want to create some web pages in that domain. How can I use the free web space available with my Virgin account?

A The simple answer is that you probably can't, although it depends a lot on what sort of free space you've been given with your internet account, and what type of URL you want to be able to quote to people. First, you'll need to arrange with the people who hold the domain records for your name to make an entry for www.yourdomain.com, or whatever the domain is called.

That's only the beginning, however. In the case of Virgin, your web space has a URL like http://freespace.virgin.net/fredbloggs/ and the same is true of many other ISPs. Unfortunately, this starts to make things difficult because all a domain entry tells a web browser is the name of the server; nothing after the first slash. Since the main purpose of a domain of your own is to have a snappy URL, having to quote a path after it rather defeats the purpose: in our example, you'd end up with www.yourdomain.com/fredbloggs.

One solution is to use an internet provider like Demon where you have a "virtual server", which means you don't need a path to identify your own files. It would be possible to open an account with Demon which entitles you to free web space on a virtual server, and then set up the domain information to point to your own virtual server. However, while this would certainly work, it's also a breach of the terms of service for your Demon account; you may find other ISPs are more lenient in that regard.

There is an alternative, although it does require the assistance of someone who operates a web server — most likely the people who host your domain name for you. This web server can be configured to redirect requests for your domain to anywhere else, effectively translating a simple server name into the name of a server, followed by the path to your web space.

If you haven't registered your name yet, an alternative is to look at one of the free services that does this more or less, such as home.ml.org, which enables you to have a name such as yourname.home.ml.org pointed to any web space on the internet. However, the way some services work means that they won't be compatible with older web browsers. It is possible to do the same thing in a way that will work with any browser, but since it involves allocating a unique internet address just for your web space, it will probably cost you money and is unlikely to be much cheaper than paying for web space to go with the domain.

In other words, there's no simple solution. If you're prepared to pay to register a domain, it's worth paying for a proper solution when you want to set up a web server to go with it.

Q I have a US Robotics Sportster Flash 33bps which used to work okay and could connect to my ISP at 33,000bps via our office phone system. I flash-upgraded to USR x2 56Kbps and all went well, but now when I connect to my ISP, it drops to only 22.6K maximum. US Robotics technical support advises me our office analogue phone network is the probable culprit and tells me to try changing my modem init string to include "S32=34" (default usually =2) to go back to 33K settings, but where is this initialisation string put?

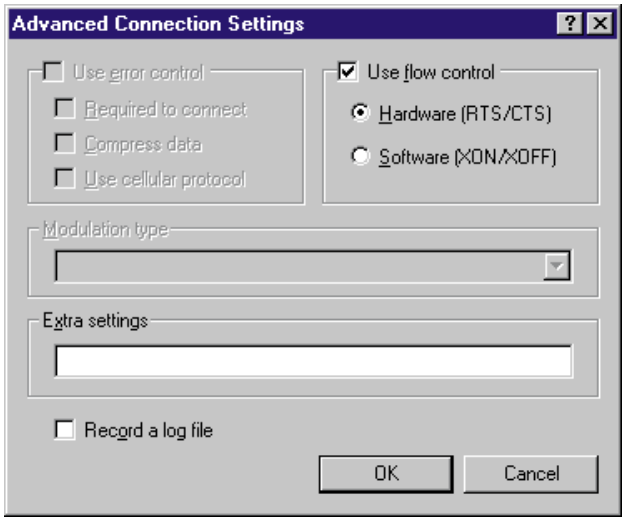

Fig 1 **It's easy to add extra settings to your modem setup**

A You're probably not alone with this problem. The claims made for 56K modems are likely to prove something of a disappointment for many people, although your case, with worse performance than before the upgrade, seems a bit extreme.

You can easily add other options to the modem initialisation string; to do it, open the control panels on your PC and click on Modems. Choose the USR modem from the list of installed modems, then click the Properties button. On the next screen, choose the Connection tab and then click on Advanced... to see the screen we've shown here (Fig 1). The extra settings can be entered in the box at the bottom of the window, and will override whatever options Windows has chosen for the modem.

Q Since I spend a fair amount of time browsing sites *ad hoc*, I like to clear out things like the Temporary Internet Files folder quite often. I have done this by opening Explorer, selecting the folder then all the files in it, and hitting the Delete key. It occurred to me to write a BAT file, with a shortcut to it on the desktop, which would do this automatically. However, when the BAT file tries to go into action, I'm just told that there are no files in that folder.

A The most likely cause of this problem is errors in your batch file. While the Windows Explorer is capable of dealing with the path to the cached files, you need to take special care with batch files as there are spaces in the path names.

The simplest solution is to check the properties of the various folders in the pathname and see what the MSDOS name is listed as, then use that in the batch file; you can simply put quote marks around long filenames, but remember to make sure that you get the names and the quotes absolutely right.

It is also worth including a line in the batch file to mark all the files as not hidden, and not read-only. In particular, Explorer may be showing hidden files that aren't visible from the MSDOS prompt. If this is the case, attrib -h filename will make a file visible to the del command: you can't use a wildcard, since MSDOS can't match it to the names that it can't see.

p246 ➢

Q I am only 15 and this is a bit of an embarrassing question. I really like this girl in my class and we often talk over the net. I can't bring myself to ask her out personally, so please could you tell me how to send her an anonymous email? My friends say it is possible but they do not know how.

A Sending an anonymous email doesn't seem to be the best way forward: would you turn up for a blind date when you didn't even know who you were supposed to be meeting? You're as likely to make someone feel they're being harassed as to entice them to meet you. However, if you don't want to use your own address to get in touch, there are solutions. You could use an anonymous remailer, which will let you hide all your personal details, but a far better solution is probably to join one of the web mail services such as HotMail at www.hotmail.com where you can choose your own address and the name that appears on messages you send.

Remember, if you're approaching someone to try and meet them, it's best to let them know who you are sooner rather than later, or it could be more embarrassing, long term.

Thousands of Real Audio files are available on the web

Q I'd like to include sound files on my web pages. What's the difference between using a format like real audio and an ordinary WAV file? Which one will give the best results?

A "Best results" is a bit of a loaded phrase. It's easier to produce a good-quality stereo WAV file for most people, as you can simply record it straight on your PC. But you're also likely to end up with a very large file if you want decent sound quality. Real audio requires a little more work to produce and can be played by anyone who has the right software, just like a WAV file. But the real advantage of the format is that listeners don't have to wait for the whole file to be downloaded before they can listen to it (sometimes). The sometimes is because for real audio to work the way it's intended, you'll have to have your files on a web server that can provide "streams", which are the way in which sound can be played without downloading it all to your system. If your ISP doesn't give you any real audio streams with your web space, there's little point in using the format as you won't get the main benefit. To the best of our knowledge, the only provider that gives you real audio streams free with web pages is Demon. Others may be able to provide it at a price.

Web systems like HotMail are better than anonymous messages

PCW Contact

Nigel Whitfield is a freelance journalist, maintainer of several internet mailing lists and consultant to a number of non-profit organisations. He can be contacted via internet@pcw.co.uk.

Getting the **message**

Nigel Whitfield has relocated and reformed. In this all-new Hands On column, he urges you to crack down on junk email, goes fishing for Perl users, and answers all your net queries.

"You have been personally selected to receive this information. Our extensive research suggests it will be of particular interest to you. But don't worry, because in spite of contacting you this time, we promise we won't be doing it ever again. Bear with us this once, however, and your life will be transformed into one of untold riches and splendour."

Internet service providers are becoming more pro-active in fighting the problem of junk mail like this. Sending such mail is prohibited by most online services, and nowadays you are far more likely to find your account terminated for doing it. You're also more likely to find yourself having problems when the internet community decides that the service provider you use is guilty of harbouring spammers.

Earlier this summer, some companies blocked mail to Enterprise customers because of junk mail that had earlier originated from that supplier. Both CIX and AOL now filter out email from known senders of junk messages, as do web-based systems such as BigFoot. And Demon's new web service allows you to kill messages without having to first download them to your system.

There are those, junk mailers among them, who simply see messages as freedom of speech. But they miss the point that we're paying for that freedom: like opening the door, being handed circulars by the postman, and then asked for money to cover excess postal charges.

The internet, although now largely commercial, still relies on co-operation for the smooth transfer of traffic from one part to another: the money you pay for your net subscription is a small contribution to one of the world's largest co-operatives. Junk mail isn't just a nuisance, it's a selfish abuse of that co-operation.

So the next time some junk mail drops into your mailbox, don't reply to it. Don't just delete it: look at the headers and see if you can work out where it came from, and if you can, complain about it. Don't stand for selfishness — and make sure your provider won't, either.

Reading web form input using Perl

Perl is one of the most popular languages for writing scripts to run on your web server; it's relatively easy to learn and most scripts can be moved easily between different systems, including Unix, NT and Windows 95 (Perl for Win32 can be found at www.activeware.com). A good place to start is the book *Learning Perl*, published by O'Reilly. If you can program in other languages, you'll probably be able to pick up a fair bit from the online manuals at

Questions & net.answers

Q I have a 33.6Kbps modem and I don't know what effect the driver has on its speed. I have experimented with the standard 28.8 driver supplied with Windows 95 and the 14.4 driver that came with the modem, and can't tell the difference. Which driver should I use?

A Always use the fastest driver you can, although you're likely to see little difference between them if you use basic drivers. The differences come where the driver tells the modem which speed to use and it's unlikely that the Standard drivers included in Windows do that, as it's one area where modems differ tremendously.

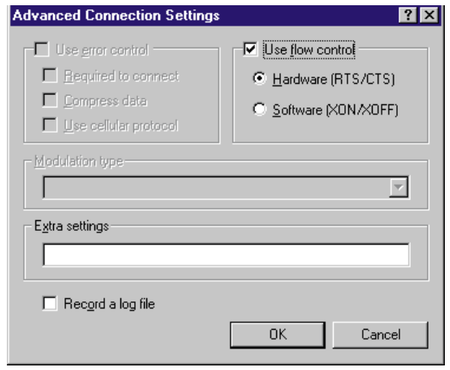

Fig 1 You can add extra modem options to override the ones that Windows selects in the driver

The best way to configure a modem is really to do it yourself, or add extra commands to the setup options to ensure the modem is being set correctly, as you can't see commands that Windows is choosing for you. For extra options, check in your modem manual for the commands to make sure it attempts a connection at the highest speed, select the modem in the control panel, choose Properties, click on Advanced and type the extra commands in the box at the bottom of the window (Fig 1). Just as important is the speed at which the computer talks to the modem. That's the "maximum speed" setting on the control panel, and for a 33.6 modem it should really be set to 115,200bps if you have a fast UART serial-port chip on your system. If you don't have a fast serial port, then the discussion is largely immaterial, as you won't get great performance whichever settings you choose.

Q I'm writing some web forms and want to use scripts to process the data that's submitted. What's the difference between the POST and GET methods and which one should I use on the pages?

A The best method to use for almost every situation is the POST method. It's more flexible, you can transfer more data, and, to an extent, it's also more secure.

The difference between the two methods is straightforward. In the GET method, the data from your form is stored in an environment variable called QUERY_STRING. The limiting factor is the size of an environment variable which varies from one system to another, depending on how much memory is allocated to the environment. In the POST method, the same information is sent as input into your script, so there's no limitation to the amount of data; the CONTENT_LENGTH variable tells the script how much there is and exactly that number of characters can be read as input.

For an example of how to decode the information sent to a script, see the section "Reading web form input using Perl" (opposite page). Remember that while the POST method is preferred, it may not be supported by all servers: for instance, the Demon Internet home pages only allow use of GET, which limits the amount of information you can have on a form.

Q I've enabled ratings in Internet Explorer to prevent my children from seeing some types of site on the internet, but when I try to browse, I find there are lots of sites that I simply can't see. Is there a way around this?

A Yes, there is a way around the problem: what you've run into is one of the most common problems with the ratings system used by some sites. Since it's a

p258 ➢

voluntary system, not all sites include the special tags that are used for rating a page. By default, Explorer won't display pages that don't have ratings, which is fine if you want to protect children but awkward for accessing some types of pages, like technical support, which are seldom rated.

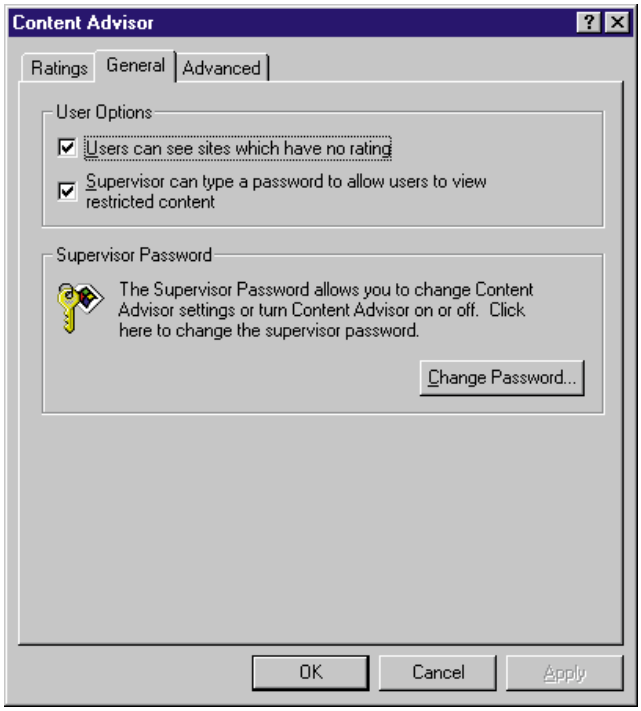

Fig 2 **You can tell Explorer that you want to see pages with no rating, at your own risk**

To change this option you should choose the ratings options and click on Settings. Choose the General tab and check the box marked "Users can see sites which have no rating" (Fig 2). Alternatively, disable the ratings system for your current setting. Either way, remember that if you don't change the options back, you'll run the risk of people being able to use the browser on your computer to access sites that you'd rather they didn't see.

Q We're getting "Content Advisor" messages when trying to access any web pages via MS Internet Explorer 3.0 (Windows 95). The ISP says this is due to a password having been entered but nobody has knowingly done it so we don't know what it is. They say it is technically possible to edit the registry and reset to "Nil" but cannot provide the way to do it. Trying to access MS help web pages are similarly restricted. Am I doomed to an internet-free life (albeit saving my marriage)? Is changing the browser the only answer?

A You can fix the problem without having to install a new browser, although a search of the Microsoft web site couldn't come up with the solution to this problem. Trial-and-error has, however, and we've noticed no ill-effects as a result. As it involves changes to the Windows Registry (which is where all the configuration information for your system is stored), you should take a backup, using a tool like the Emergency Recovery Utility.

The program you need to make the changes is called Regedit; you can start by clicking on the Start button, choosing Run and then typing regedit into the box. The Registry Editor is similar to Windows Explorer: to find the information you need to change, open folders in the left hand pane, starting with HKEY_LOCAL_MACHINE, and navigate your way down to

`HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Policies\Ratings`

In the right pane, you'll see a display similar to the one in Fig 3: the hexadecimal codes are the password, encrypted so that it can't be broken. You don't need to worry about what it was set to; with the Ratings folder highlighted, press the Delete key and the whole folder will disappear.

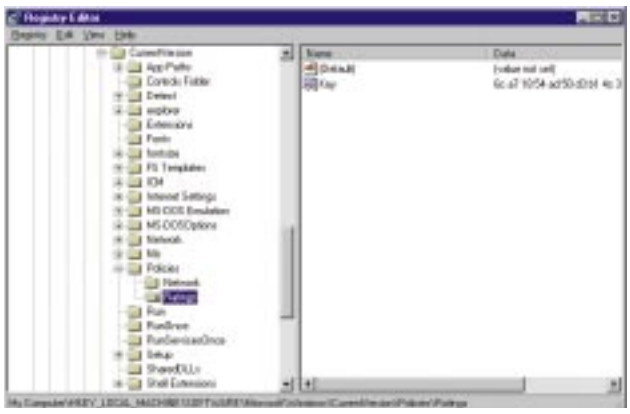

Fig 3 **Regedit will solve your problem, but incorrect use is risky**

Next time you go into Internet Explorer, from Options, choose Security. You'll notice that the system still thinks that ratings are enabled. Click on the Disable Ratings button and you'll be asked to enter a password. You can now use a new password of your choice and you'll then be able to turn the content advisor off and on at will. If you wish to leave it turned off without a password set, simply enter a password, then follow the procedure above to delete the registry entry again.

Q Is there a way that I can run a mailing list using a standard dialup account from my internet provider?

A In theory, you can run a list. There are versions of list software for machines running Windows and Mac operating systems. You can manage a list by hand, simply by maintaining an alias list in a mail program like Pegasus mail. However, bear in mind that if you want everything automated, you'll need to be able to identify messages for the list automatically, which may not be easy if you only have one mailbox from your provider.

It can be time-consuming to distribute messages to a mailing list, depending on how your provider's email system is set up. It may want to check each address as it's sent out. For some addresses, that could take as long as a minute online. With a large list, that time can soon mount up.

The best solution is to find someone to host a list for you, on a system that's permanently connected to the net. Look for a system with software like Majordomo which can be remotely controlled by sending email from your home account (and via the web) so you can do admin without running up your own phone bill.

If you decide to run a list of your own, invest in an account with a provider that lets you have multiple mailboxes so you can allocate a unique address to the list, making everything easier to manage.

p260 ➢

www.perl.com too.

Scripts that create pages on-the-fly are easy: all you need is Print commands to create the HTML you want and you can turn data on disk into nicely formatted pages. But to be really useful, you'll need to read information from the web server and to store it in variables within your script. Some servers provide tools especially for this, but a few lines of Perl can do the trick for every web form.

First, you'll need to use the POST method of submitting data: that means your web page will need a line something like

```
<FORM METHOD='POST' ACTION=
'myscript.cgi'>
```

In your script, you'll need to read the appropriate number of characters from standard input, which effectively means that it behaves as if the form data was typed into the script. The web server uses the environment variable CONTENT_LENGTH to say how many characters there are, and URL encodes the actual data entered into the web page. That means that all the fields from your form are combined into one line, with certain characters represented by a % followed by the hexadecimal ASCII code. For example, a form with a field called "name", into which "Nigel" is entered and a message field with "Hello world", would be encoded as:

```
name=Nigel&message=Hello%20world
```

To make use of the information, you'll need to split the line into parts and change encoded characters back to their original form. The Perl for this is shown above.

Perl does the job

```
1       read(STDIN,$qs,$ENV{'CONTENT_LENGTH'}) ;
2       foreach ( split('&',$qs)) {
3               $_ =~ s/\+/ /g ;
4                       $_ =~ s/%(..)/pack('c',hex($1))/ge ;
5                       ($key, $val) =  split( '=', $_ , 2 ) ;
6                       $i{$key} = $val ;
        }
```

1. This line reads from standard input into the variable $qs, the exact number of characters specified by the CONTENT_LENGTH environment variable.
2. Now we split the query string ($qs) into parts for each field; they're separated by the ampersand (&) character. Since we've not specified a variable in the "foreach" statement, $_ will take on each value in turn. For our example data, it would first be name=Nigel.
3. A search and replace; the first string (+) is replaced by a space, wherever it occurs.
4. This line looks for a % symbol followed by two characters: the brackets around the two periods mean that the characters will be saved as $1; the pack command is used to turn the hex code back into a single ASCII character. Once that's done, everything's decoded.
5. Now we chop up the information at the equals sign: to the left is the name of the field, and to the right is the info it contains.
6. Finally, we use Perl's associative arrays to store the data: by doing this, we don't have to alter the code for different forms.

The rest of your script can do whatever it likes after you have grabbed the input from the form: writing it to a disk file or creating a page, for instance. When you want to refer to data from the web page, you can do it simply by referring to, for example, $i{'name'}. And the rest of your script? That's up to you…

■ Starting this month, *Cutting Edge Net Answers* becomes *Hands On Internet*, with more comment and hands-on material. If there are any relevant subjects you'd like to see covered, please let us know.

PCW Contact

Nigel Whitfield is a freelance journalist, maintainer of several internet mailing lists and consultant to a number of non-profit organisations. He can be contacted via internet@pcw.co.uk.