The 'Hosts' File

Description: Steve and Leo reveal and describe the HOSTS file which is hidden away within every Internet-capable machine. They explain how, because it is always the first place a machine looks for the IP address associated with any other machine name, it can be used to easily and conveniently intercept your computer's silent communication with any questionable web sites you'd rather have it not talking to.

High quality (64 kbps) mp3 audio file URL: http://media.GRC.com/sn/SN-045.mp3
Quarter size (16 kbps) mp3 audio file URL: http://media.GRC.com/sn/sn-045-lq.mp3

Leo Laporte: Bandwidth for Security Now! is provided by AOL Radio at AOL.com/podcasting.

This is Security Now! with Steve Gibson, Episode 45 for June 22, 2006: The Hosts File. Security Now! is brought to you by Astaro, makers of the Security Gateway, on the web at www.astaro.com.

Well, now I'm feeling guilty. I pressed Steve into service a day early. I'm taking a little much-needed break this week, off at the beach house.

Steve Gibson: It's funny because you said you were on vacation. I said, "Vacation from what?"

Leo: Like I don't do nothing. Oh, yeah, you working stiffs don't understand, I have a complicated life.

Steve: I guess those of us who don't have a boss are sort of, you know, actually we're never on vacation because we...

Leo: Well, I was going to say, you don't get – do you ever take vacations?

Steve: No, I don't.

Leo: You just don't even need a vacation.

Steve: I'm just bored. I love what I'm doing. And, you know, if ever I go somewhere, it's like I'm just spending all my time thinking, hmm, just to think all the things I could be getting done if I was still home.

Leo: Here's a man who loves his work, doesn't even consider it work.

Steve: I really do.

Leo: Yeah. Well, when you came up to Canada, that was kind of a two-day vacation. I can't get him on a geek cruise, folks. I've tried and tried, and he says, "I'm not going on any boat."

Steve: Actually, I really liked the two-day Canada turns because I had a whole day going, you know, across the country to the other side of the world to read and to do research and things that normally I just don't give myself the opportunity to. So it was like little forced breaks.

Leo: See? See? See?

Steve: Yeah. It was good.

Leo: Actually, it's one of the things I like to go to Canada, same reasons I like to go to Canada. I pretend that it's the other side of the world, and I tell people, oh, I won't be able to do anything for the whole next week because I'm going to be out of the country. Of course it's not really out of the country exactly. It's pretty easy to reach me in Canada. But I pretend. Uh-oh, now my secret's out. I'm going to have to get a job in Alaska. I'm going to have to get a job in Timbuktu.

So today we're going to talk about something that people – everybody has. Everybody's got one. Many people don't know they have one.

Steve: It's free.

Leo: It's free. It's on your system.

Steve: It can be very effective.

Leo: It can be a very useful tool in fighting bad guys. We're talking about the hosts file.

Steve: Yeah.

Leo: What is the hosts file?

Steve: Well, it's interesting. It's been around from the dawn of TCP. As far as I know, the hosts

file, as it's called, h-o-s-t-s, is present in any system which is connecting to the Internet that has a TCP/IP stack. That is, you know, the standard Internet protocol glue that we all use for hooking things up.

Leo: UNIX uses it, Linux uses it, Windows uses it, Mac uses it, OS 2 used it. I mean, it's pretty universal.

Steve: Yeah. And essentially what it is, it's sort of like a first DNS lookup which any TCP-equipped system will reference – and here's the key – before it goes and uses the official DNS system to map a domain name to an IP address.

Leo: Oh, interesting. So when you type in www.grc.com, normally, I mean, we normally think of it as going out to the DNS server of your Internet host, or maybe of the main domain servers, and saying, well, where is GRC.com? But you're saying it checks this hosts file before it even does that.

Steve: Well, a few things happen. When you enter a domain name into your browser, systems are now maintaining a local cache of these names. So, for example, it knows it just recently looked up, one way or another, it looked up the address for a domain name, and it doesn't have to go and make another query because it would be crazy to have your computer constantly asking your ISPs or whatever DNS server you had configured what the IP address was for a domain name. So certainly it makes sense to cache those locally, that is, keep a record of them. In fact, these are timed so that the system knows how long it's been since it last updated its own local cache, so that it can, if the IP address were to change, it would know that, hey, I've had this for three days, I ought to make sure that this is still the same IP address.

Leo: Now, actually there have been problems with the – I remember the XP cache because it didn't handle it properly, didn't flush it properly. So does the cache persist across a reboot, or...

Steve: No, it won't persist across a reboot. There have been problems, in fact, one of the annoyances is called "negative caching." And what that means is that, if your computer asks a remote server for the IP address of a domain name or a remote machine, and for whatever reason the request fails, the system will – it's called "negative cache." It'll say there is no IP for this domain. So it essentially remembers a negative. And that's a problem because you might have a momentary glitch somewhere, or something on the Internet might have failed that would cause you to get a false negative domain name. And then your system won't go out and try it again. It thinks it knows better. There actually is a command that you're able to execute from the command line to flush your local DNS cache. And in some cases doing that restores connectivity. But it's nuts that that's what you have to do.

Leo: I thought they'd fixed that, but – so it wasn't a bug. It was actually – or was it a bug?

Steve: It's, oh, it's a feature.

Leo: It's a feature.

Steve: That's not a bug, it's a feature. So anyway, the idea is that, you know, there's a lot of

plumbing going on in just in DNS on our own systems. The really interesting thing is that, prior to any of this remote access, a simple text file is checked that exists on all of our systems. I mean, it's even there by default. It's installed when you install Windows. Most people never, who are not into all this kind of stuff, never go any further, never pursue it, never even know about it. But it's there. It's under your Windows directory. That might be, you know, WinNT or Windows, depending on whatever your system calls it. Then all of our systems in Windows have a subdirectory called "system32." Under that directory, strangely enough, it's under a subdirectory called "drivers," and then "etc," that is to say, e-t-c. So it's even in an obscure location.

Leo: The etc location is actually where UNIX and Linux stores it, in /etc.

Steve: Right. And in fact it makes sense, I mean, it's sort of like, you know, extra little stuff to go along with TCP, or rather with the IP protocol. So what this is, this text file – and anyone listening to this who doesn't know about this who's curious can look at their own hosts file under their system32/drivers/etc, if they're in Windows...

Leo: Of course Windows is going to warn you, well, you shouldn't be looking in here, what are you looking in here for, don't be going in there.

Steve: Oh, yeah.

Leo: And all sorts of other stuff.

Steve: Yeah.

Leo: But just so you know that's what's going on. It's okay.

Steve: All it is, is it's a simple table that is very much like DNS. It matches IP addresses to domain names. So, for example, a very commonly used name is "localhost." Localhost, by sort of IP gods, is used to refer to your own machine that also has an IP address of 127.0.0.1. We've never talked about this before, and we probably ought to spend some time at some point talking about some of the, you know, these things related to the localhost and the local network. But 127, the IP address 127.0.0.1 is in all IP stacks a self-reference. It's another way of talking to you about your own local IP. So the hosts file will have an entry there that just says 127.0.0.1 is the same as localhost. So if you use the word "localhost" just like a domain name, that relationship between localhost and 127.0.0.1 comes from the hosts file. And the way this was used initially was within a LAN, that is, within a local network this is a nice way of associating machine names and their IPs.

For example, I use it myself, or actually I did in the past. Now I have full-blown DNS locally. But I used to have all of my machines named. And then they had fixed IP addresses. So I had a hosts file which I installed in all of these machines – and by "installed" I mean just, you know, copied one hosts file to all of them. And that allowed me to refer to the machines, for example, for filesharing and things, not needing to use IP addresses, not even needing to use Windows browser technology, which can be sort of flaky and has some version dependencies, but it would use the hosts file. Basically I was providing my machines with names and allowing them to determine that the IP addresses were for each of the machines in my network. So it's a cool thing. Now...

Leo: So you don't need to mess with this. It's handled kind of automatically. In fact, by default it only has one line, which is the localhost. But you decided to use it.

Steve: Well, I decided to use it. And it works great for that purpose.

Leo: Do you recommend people do that?

Steve: Well, I mean, I wouldn't recommend someone who's not into networking and stuff to do it. But doing it really does work. And in fact, if someone gets into it more, it turns out that there's a way you can put an entry in a hosts file to refer to another hosts file, sort of an include facility, where you say include a file from somewhere else in my hosts file. What that allows you to do is actually have a single master hosts file that all the machines in your network will automatically pull theirs from. So I didn't want to get too complicated here. But in fact I had one copy of a hosts file on my network. All the machines knew where to get theirs from, that is, in their hosts file I had a reference to this, sort of an included file. And so as I added or moved machines around, I made one change. And what was convenient was it is a really robust solution for machines.

Now, again, the problem is, most people are using DHCP behind their own NAT routers. So the IP addresses of their machines may be floating around. That's one thing I'm not doing in my network. I don't use DHCP. All of my machines are known IPs. So this is only really applicable in a case like that. But the reason we're really talking about this is from a security and a privacy standpoint because very clever people have figured out how the hosts file can be used to prevent machines from accessing domains you don't want them to. And, I mean, having said that, a lot of people who are listening to this are now going, oh, I know how to do that. I mean, essentially you edit your hosts file to – for example, I'm just looking at mine right now. And for example, I've got an entry that says 127.0.0.1.

Leo: That's your localhost. That's your local machine.

Steve: My localhost. And then tab, and it says fastclick.net, or mediaplex.com, or focallink...

Leo: So these are online advertising sites.

Steve: These are, exactly, these are the people we would rather not have our computer contact. For whatever reason, we don't want cookies, we don't want ads, we don't want our computer for any reason to go touch doubleclick.net or linkexchange.com or valueclick.com. I mean, I've got a whole list of these. And the beauty of this is it is solid and robust. And my machine, anytime I go to a site that has references to those domains, the browser is going to say, oh, what's the IP, so it can go get an ad or get whatever it, you know, who knows, a web bug or spyware or anything that it might be pulling from those domains. The first thing the computer does is look up in the hosts file to see if that domain has been defined. If so, that's the address it uses, in other words, 127.0.0.1, meaning itself. So that request goes nowhere. It's immediately canceled, and the browser goes about trying to do, you know, load the rest of its web page. So it's a very clever, simple, and very efficient means for preventing your machine from ever going elsewhere, or, I mean, to an expanding, evolving, or growing list of, like, bad or banned domains.

Leo: Now, just as you've used it to block bad guys, bad guys also use it. There is spyware that modifies the hosts file.

Steve: Yes.

Leo: Mean.

Steve: Of course, we could expect that anything valuable like this that could be used for good would also be grabbed and used for bad. You're able to – there are ways you can protect your hosts file from modification. For example, you could write-protect it or remove write privileges from the file so that only the administrator of your system, which you should not normally be running under, would be able to modify the hosts file. Then, if some spyware got into your system that was not running as an administrator, it would not be able to change the file. Sometimes just clicking the write-protection flag is enough to defeat these things.

We should say, though, that what has evolved is sort of an interesting – it's not really a subculture, but it's a bunch of people who manage public hosts files. That is, there are lists of hosts files on the 'Net that you can download that are being maintained by people and contain a huge directory of places you probably don't want your computer to go. That's where I got my list from some time ago. And those domain names don't tend to be changing very often. So if anyone was curious about this approach, they could just put "hosts file" into Google. I think you mentioned that Wikipedia has a great discussion about the hosts file.

Leo: They do, yeah, yeah. And if you use Spybot Search & Destroy, or I'm sure other spyware programs, it also will do that. In fact, there's a command in Spybot that just says "populate my hosts file," and it will add a whole, you know, changed hosts file that blocks all of those well-known sites. Really hundreds of names. And I think it then locks it down, as well.

Steve: It would make sense that it would because, again, there have been exploits that modify the hosts file. On the other hand, remember that, once something bad gets into your system, you're pretty much hosed anyway. I mean, you can, you know, you can never really know what was done to your machine. So it's certainly the case that something nasty getting into your system could modify your hosts file. On the other hand, it could modify anything that it wanted to if it's able to modify your hosts file. The power of the hosts file is not so much that it's an absolutely perfect solution, but that it's – I mean, because none exist in the Windows environment, or in fact in any but an ultra secure system, which sort of becomes no fun to use, depending upon how tight the security is. The beauty of it is that it is an extremely reliable, very efficient means to prevent a computer from going places you don't want it to. And it's simple to manage and maintain. Its format is just IP address and then a tab character or a space or something. Mine's all lined up nicely because I have 127.0.0.1, tab, and then the domain name. So it's very easy to add or remove items. And it just simply prevents your system from asking the Internet for the actual IP address of any domains that are listed there. So you can just neuter these things. And, I mean, it's a tremendous little tip.

Leo: One little word of warning. If you are running a web server on your local system, as I do sometimes for test purposes, mapping to localhost will have an unpredictable result because your web server may attempt to fulfill that request. So in that case you might want to use 0.0.0.0, or I understand you can just use 0, and it'll still have the same effect of...

Steve: Yeah, it'll just, exactly, just goes nowhere.

Leo: I don't know where it's at.

Steve: It just stops the – it stops the search. It nips it in the bud, saying there's no reason I want anybody, you know, I ever want this system going to DoubleClick. And it just – it can't get there because it's given a bogus IP address.

Leo: I think Windows Defender also will monitor your host files. That's the new Microsoft anti-spyware program. And I'm just looking at the Wikipedia article, and they mentioned that that was how Mydoom worked, to block your access to Symantec and other antivirus sites. You may remember, if you got that bad virus, you couldn't then do any research on how to get rid of it because you couldn't get to any antivirus sites. And that's exactly how it did it, it just modified the hosts file.

Steve: It just puts a bunch of entries in there for the anti-spyware and antiviral vendors and prevents your computer from going there, too. And in fact, you know, it's probably worth having people take a look at the hosts file, just to make sure that it still looks benign.

Leo: Absolutely. That would be a sign of tampering, wouldn't it.

Steve: Yeah.

Leo: Yeah. All right. Well, there it is, everything you ever wanted to know about hosts files. And if you want to know more, we'll put a link to the Wikipedia article in our show notes. But you might also check there because there are additional links in there to a lot of other good resources, including somebody who says, you know, it's a myth that changing your hosts file will protect you from malware. Obviously it won't, but it'll help, you know, it'll help reduce your impact of malware.

Steve: Well, yes. And again, I want to make that point because I think it's a good one. It's that the hosts file being set up to block a whole bunch of malicious domains can prevent you from getting that malware in the first place. So, and it's – and a non-compromised system. I mean, there are a lot of places that are not malware, they're just, you know, they're heavy-duty marketing, Internet marketing companies that are using cookies and technologies to track you around the 'Net that you'd rather not have any contact with. They're just, you know, sort of gray zones. They're not black, they're just gray. So the hosts file just prevents your machine from touching those. And again, it can be abused. But, you know, so can everything.

Leo: Right. It's a process. There's no one silver bullet in security. I wish there were.

Steve: Yeah, I mean, it's sort of like the argument we touched on when we were talking about open ports about, you know, there are people who say that stealthing your system doesn't provide you any security. It's like, well, okay, I mean, I've heard the argument a lot. It's better to be invisible, I think. Why not be? It doesn't cost anything to be invisible.

Leo: Right. As long as you just don't fool yourself that you've done everything you can do at that point.

Steve: Well, or that everything that you've done is going to make you absolutely invincible because...

Leo: Right, because there's no such thing.

Steve: ...you know, that we can't achieve. I'll also mention that my own Security Now! page, I'm going to put a whole bunch of hosts file references on that page of stuff that I use and that I have found over time. So anybody who's interested in pursuing this independently can not only check Wikipedia and the link you're going to have, Leo, but it's been a long time actually since I've put together some notes for a show because most of what we've been doing hasn't involved off-show resources. But here, in the case of the hosts file, I think we really have – there's just so much that's been done. It's a tremendous resource.

Leo: And once you understand it, I mean, I think this is something people don't even know they have. So once you know you've got it, this is worth kind of learning more about. And we'll put all that in the show notes. Of course show notes are at GRC.com/securitynow. That's where you'll also find a transcript of this and every show, thanks to Elaine, our wonderful transcriptionist, and 16KB versions for the bandwidth-impaired. You know, you've started something, Steve, because I've started doing 16KB versions of TWiT, as well, and my radio shows, because people like – if you're on dialup, people like to be able to download it and still listen, even though...

Steve: I'm really glad to hear that, Leo. I mean, I know how popular they are.

Leo: They are.

Steve: Oh, they're incredibly popular. I mean, I think maybe it's easy to be a bit of a bandwidth snob and to say, oh, well, you don't have cable, you don't have DSL? But there's a lot of people still using a modem.

Leo: You bet. And I'm not so much a bandwidth snob as an audio snob. I listen to those 16KB versions, and I go, ooh, gosh.

Steve: Oh, I know.

Leo: Ooh.

Steve: I know.

Leo: But you get the content, and that's the most important thing.

Steve: Yeah.

Leo: So we do put those up. We also thank our great friends at Astaro for providing the support, the sponsorship for this podcast. We mentioned a couple of weeks ago the Astaro Command Center, v1. It's out now, free for users of Astaro's great Security Gateway. If you're a network administrator you should see this thing. It is a great way to manage and control multiple gateways from just one dashboard, and a beautiful world map, and, oh, I mean, it just looks great. It's free in the Products section at Astaro.com. And of course Astaro is a great solution for any small or medium business network that needs superior protection from spam, from viruses, from hackers, as well as complete VPN capabilities, intrusion protection, content filtering, and an industrial-strength firewall, in an easy-to-use, high-performance appliance. Contact them at 1-877-4AS-TARO, or, again, Astaro.com. And don't forget that there is a solution for home users, too, that's absolutely free. Let's see, what else? I guess that's it. I guess...

Steve: Security Now!, and...

Leo: Security Now! – oh, I forgot the most important thing. SpinRite, ladies and gentlemen. Yeah, Steve never takes a vacation. That's because his product is always being improved. SpinRite is free – is not free but is available from GRC.com. It is SpinRite, and you can read more about all the testimonials on SpinRite at SpinRite.info. I think anybody who has a hard drive – I think that means everybody – should have a copy of SpinRite, the ultimate disk recovery and maintenance utility, from SpinRite.info. Have I mentioned everybody? I think so.

Steve: I think it's covered.

Leo: The time has come, I hate to say it, to say goodbye. But we'll see you next week. Thank you, Steve.

Steve: Absolutely. And in the meantime, people can check out their hosts file. It's just such a neat little perfect solution for keeping your machine where you don't want it to go.

Leo: Yes. Yes. You're a good host. Take care, Steve. We'll see you next Thursday.

Steve: Right.

Leo: On Security Now!.