

CITS4407 Open Source Tools and Scripting

Network tools

Unit coordinator: Arran Stewart

Overview

This week:

- communicating over a network

Note that this lecture gives a *simplified* explanation of networking concepts.

For a fuller one, you should take a look at the materials for a unit like CITS3002 “Networks and Security”.

Commands that use the network

We've already seen a number of commands that contact other systems over the Internet.

- Whenever we run `apt-get update` the `apt-get` command contacts computers holding copies of Ubuntu's software repositories, and fetches the current list of available packages.
- When we run `apt-get install`, the `apt-get` command download copies of the software packages we want.
- `git clone`, when we give it a URL like <https://github.com/cits4407/assignment1.git>, contacts GitHub's servers to fetch a copy of the repository we're cloning.
- The command `wget` (used in labs) retrieves copies of files or pages on Web servers.
- Whenever we visit a website in a web browser, copies of the pages we see are being fetched from remote servers.

Hosts, domains and IP addresses

Everyone should be familiar with what a web address looks like:

<https://cits4407.github.io/resources/>

A web address is formally known as a *URL* (Uniform Resource Locator).

URLs

A URL contains the following parts:

- `cits4407.github.io` – this specifies a *host*: a system that can be contacted over the Internet. We can think of `cits4407.github.io` as being an address, that tells us how to get to that host.
- `.github.io` – this specifies a *domain*: a collection of hosts. (We could think of this as being like a suburb or region, for an address.)
- `/resources/` – this specifies the location of a particular file or directory on the host.
- `https://` – this specifies a particular **protocol** – like a common language spoken by your computer and the remote one – which should be used when contacting the host and asking for a particular file or directory.

IP addresses

In fact, hostnames like `cits4407.github.io` are just a *human-readable* way of referring to a computer on the Internet.

Your computer's operating system knows how to convert a human-readable hostname like `cits4407.github.io` into an *IP address* – an address that is convenient for *computers* to use when locating a remote host.

They look like this:

- `172.217.25.132` (an Internet Protocol Version 4 address)

or this:

- `fe80::7ef3:e931:830c:6471` (an Internet Protocol Version 6 address)

Tools for working with remote hosts – ping

It can often be helpful to know how to diagnose network problems using Linux tools.

One of the most fundamental tools is `ping`.

It sends a very small network message to a remote host, and asks that host to respond.

```
ping www.google.com
```

will try and contact the host `www.google.com`.

ping

If the remote host is contactable, ping will produce a report like this:

```
arran@barkley:~$ ping www.google.com
PING www.google.com (172.217.25.132) 56(84) bytes of data.
64 bytes from syd15s03-in-f4.1e100.net (172.217.25.132): icmp_seq=1 ttl=113
time=48.2 ms
64 bytes from syd15s03-in-f4.1e100.net (172.217.25.132): icmp_seq=2 ttl=113
time=48.1 ms
64 bytes from syd15s03-in-f4.1e100.net (172.217.25.132): icmp_seq=3 ttl=113
time=48.4 ms
^C
--- www.google.com ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2002ms
rtt min/avg/max/mdev = 48.156/48.278/48.471/0.226 ms
```


Just as web browsers use a particular protocol or “language” to talk to remote hosts, so does ping – it uses ICMP (Internet Control Message Protocol).

You may not *always* be able to use ping and ICMP to contact a remote host, even if it is “up”: sometimes, network administrators will block the use of ICMP, on the grounds that only malicious hackers would want to know whether a remote host is up.

traceroute

The traceroute command reports whether a remote host could be contacted – but it also lists all the “hops” taken by network traffic as it goes from our local system to a remote host.

```
arran@beaker:~$ traceroute www.google.com
traceroute to www.google.com (142.250.70.228), 30 hops max, 60 byte packets
 1  kermit.lan (192.168.10.1)  0.285 ms  0.337 ms  0.429 ms
 2  10.20.26.6 (10.20.26.6)  7.504 ms  8.451 ms  8.422 ms
 3  203.29.134-193.tpgi.com.au (203.29.134.193)  9.298 ms  9.269 ms
10.089 ms
 4  nme-sot-dry-crt1-Be40.tpgi.com.au (203.219.107.225)  50.808 ms
50.779 ms  51.714 ms
 5  27-32-160-69.static.tpgi.com.au (27.32.160.69)
50.762 ms 27-32-160-5.static.tpgi.com.au (27.32.160.5)  51.736 ms
51.695 ms
 6  72.14.213.0 (72.14.213.0)  53.631 ms  51.545 ms  52.374 ms
 7  * * *
 8  mel05s02-in-f4.1e100.net (142.250.70.228)
47.809 ms 216.239.54.50 (216.239.54.50)  48.874 ms
...
```

Diagnosing network problems

If you are unable to run commands that make use of the network (`git clone`, `wget` and so on) – Service Desk or teaching staff will often ask you to run a command like `ping` so they can tell whether the problem is with the command (maybe `git` has a bug) or your computer.

localhost

Unix-like systems use the name `localhost` to refer to the local system.

You can ping the localhost:

```
arran@barkley:lectures$ ping localhost
PING localhost (127.0.0.1) 56(84) bytes of data.
64 bytes from localhost (127.0.0.1): icmp_seq=1 ttl=64 time=0.066 ms
64 bytes from localhost (127.0.0.1): icmp_seq=2 ttl=64 time=0.069 ms
^C
--- localhost ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1004ms
rtt min/avg/max/mdev = 0.066/0.067/0.069/0.008 ms
```

localhost

However, the `ping` command in this case isn't *really* using the network.

Messages sent to `localhost` never actually leave your computer.

Retrieving a web page

We will examine what happens when you visit a web page like <https://cits4407.github.io/resources/>.

On Linux, a good way to see what is going on is to run

```
curl -v https://cits4407.github.io/resources/ 2>&1 | less
```

Retrieving a web page

```
curl -v https://cits4407.github.io/resources/ 2>&1 | less
```

% Total	% Received	% Xferd	Average	Speed	Time	Time	Time
Current			Dload	Upload	Total	Spent	Left

Speed

^M	0	0	0	0	0	0	0
----	---	---	---	---	---	---	---

0 --:--:-- --:--:-- --:--:-- 0* Trying 185.199.109.153...

* Connected to cits4407.github.io (185.199.109.153) port 443 (#0)

* found 129 certificates in /etc/ssl/certs/ca-certificates.crt

* found 521 certificates in /etc/ssl/certs

* ALPN, offering http/1.1

* SSL connection using TLS1.2 / ECDHE_RSA_AES_128_GCM_SHA256

* server certificate verification OK

* server certificate status verification SKIPPED

* common name: www.github.com (matched)

* server certificate expiration date OK

* server certificate activation date OK

* certificate public key: RSA

* certificate version: #3

* subject: C=US,ST=California,L=San Francisco,O=GitHub\, Inc.,CN=www.github

* start date: Wed, 06 May 2020 00:00:00 GMT