NfSen top talkers

Introduction

Goals

- Use NfSen to find out which hosts are generating the most inbound and outbound traffic on your network

Assumptions

Your router is sending netflow records to your srv1 shared Server, and that host is running NfSen to collect this data. I

http://oob.srv1.campusX.ws.nsrc.org/nfsen/nfsen.php (http://oob.srv1.campusX.ws.nsrc.org/nfsen/nfsen.php)

Generate some traffic

Firstly, we need to generate some traffic passing through your routers. One one of your campus hosts do the following:

```
$ cd /tmp
$ wget https://nsrc.org
$ cat index.html
```

This will download a moderate file of about 100KB. Repeat this a couple of times more. It will take around 5 minutes before this shows as a spike in NfSen.

> *Note that 100KB is 800,000 bits, which when averaged over 5 minutes (300 seconds) is about 2.7 kilobits per second. Not a big spike, but we should be able to find it.*

Exploring flow records

Now let's use NfSen to explore the traffic flows in the network, with the aim of finding out who was been downloading the most data. Look carefully at the output generated at each step - ask an instructor to explain if you don't understand what you see.

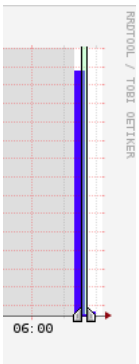
Navigate to Detail page

The NfSen home page shows a matrix of graphs: flows per second on the left, packets per second in the middle, bits per second on the right. Click on the top-right graph (bits per second, one day view) to get to the Detail page.

Select time window

Change from "Single Timeslot" to "Time Window":

Once you have done this, the vertical selector arrow and line in the graph window can be split.



Pull the left half of the arrow to the left and the right half to the right, to select the time period of interest. Then you should see some summary statistics appear in the table below the graph, for the time period you have selected:

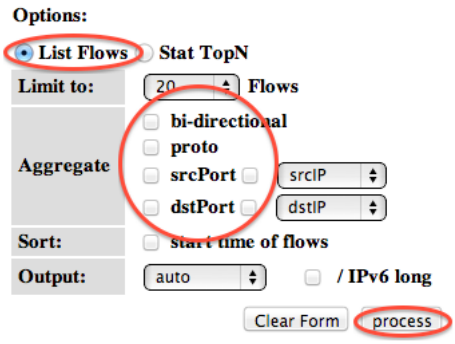
Statistics timeslot Jul 17 2013 - 20:50 - Jul 17 2013 - 21:00

Channel:	Flows:					Packets:					Traffic:				
	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:
bdr1	4.7 /s	1.0 /s	3.7 /s	0.0 /s	0 /s	110.1 /s	105.3 /s	4.4 /s	0.3 /s	0 /s	313.0 kb/s	309.6 kb/s	3.1 kb/s	254.0 b/s	0 b/s
	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:	all:	tcp:	udp:	icmp:	other:
TOTAL	4.7 /s	1.0 /s	3.7 /s	0.0 /s	0 /s	110.1 /s	105.3 /s	4.4 /s	0.3 /s	0 /s	313.0 kb/s	309.6 kb/s	3.1 kb/s	254.0 b/s	0 b/s

Summary statistics

List individual flows

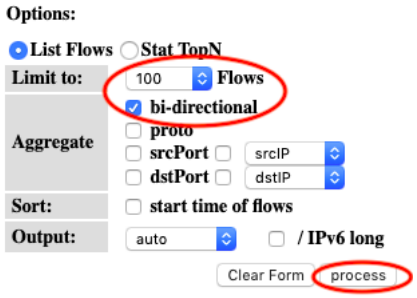
Select “List Flows”, make sure none of the “Aggregate” boxes are checked, and then click `process` . This will display some flows at the beginning of the time period.



List flows

Increase the limit from 20 flows to 100 flows. Notice that much network traffic consists of large numbers of very small flows - for example a DNS query/response will be two flows, one from client to DNS server, and one back again.

By selecting “bi-directional” you can get NfSen to associate the inbound and outbound flows into a single line:

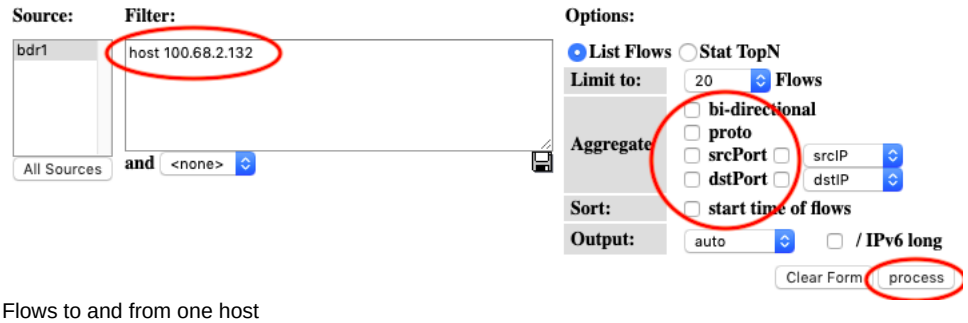


Bi-directional flows

However it’s still too much work to wade through this looking for interesting traffic. Uncheck the “Bi-directional” box before continuing.

Flows to/from one host

If we know which host we want to examine, we can apply a filter to show only those flows to and from that host. Do this by entering “host 100.68.X.Y” in the filter box, and then pressing `process` again. (Replace 100.68.X.Y with the address of one of a host on your campus. e.g. 100.68.1.131)

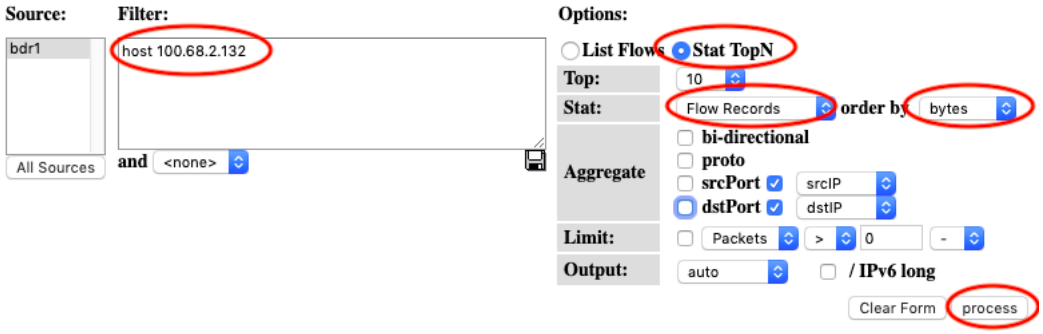


Flows to and from one host

This is a little better, but we would still have to wade through lots of small flows to find anything significant. We need to take a different approach.

Largest flows

The next thing we can do is to get NfSen to sort the flows by number of bytes. Remove any filter from the Filter box; select "Stat TopN", stat "Flow Records", order by "Bytes". Ensure all the aggregate boxes are all unchecked, then press `process`



Find top flows by bytes

```
** nfdump -M /var/nfsen/profiles-data/live/gw  -T  -r 2018/02/21/nfcapd.201802211020 -n 10 -s record/bytes
nfdump filter:
any
Aggregated flows 13604
Top 10 flows ordered by bytes:
Date first seen          Duration Proto      Src IP Addr:Port           Dst IP Addr:Port   Packets    Bytes Flows
2018-02-21 10:21:06.592   107.835 TCP     128.223.157.25:443    ->       100.64.2.6:51662      3549  5.6 M     1
2018-02-21 10:16:54.811    60.404 UDP     172.217.25.163:443    ->      100.64.2.13:62971      1521  1.8 M     1
2018-02-21 10:15:40.522   104.279 TCP      23.77.131.134:443    ->      100.64.2.23:54716      1187  1.7 M     1
2018-02-21 10:14:54.173   188.810 TCP     216.58.199.46:443     ->      100.64.2.18:41538       794  1.2 M     1
2018-02-21 10:17:06.475    34.461 UDP     172.217.25.174:443    ->      100.64.2.13:57798       529  692909    1
2018-02-21 10:15:26.127   290.132 UDP     100.68.100.254:35496  ->   100.68.100.250:9996        931  437276    1
2018-02-21 10:15:37.882   106.949 TCP      23.215.58.107:443    ->      100.64.2.23:54705       323  372373    1
2018-02-21 10:22:21.925     7.922 TCP       100.68.2.130:80     ->      100.64.2.30:50474       147  318716    1
2018-02-21 10:15:15.717    26.366 UDP     172.217.25.163:443    ->      100.64.2.16:58803       288  316884    1
2018-02-21 10:18:13.713     3.210 TCP      93.93.130.214:443    ->      100.64.2.23:54905       190  300034    1
Summary: total flows: 13704, total bytes: 54234573, total packets: 108855, avg bps: 710274, avg pps: 178, avg bpp:
```

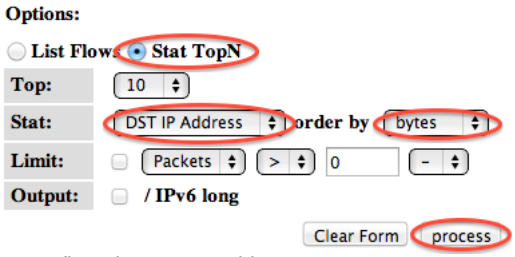
Output: top flows by bytes

This is a definite improvement, as the flows with the largest number of bytes are shown first. However there's a problem - we are still looking at individual flows. It's possible that many small flows to the same host would add up to a large amount of traffic, but we wouldn't see them at the top of this list.

Inbound traffic grouped by receiver IP address

What we want to see is a single line for each host in our network, showing the total amount of traffic delivered to that host.

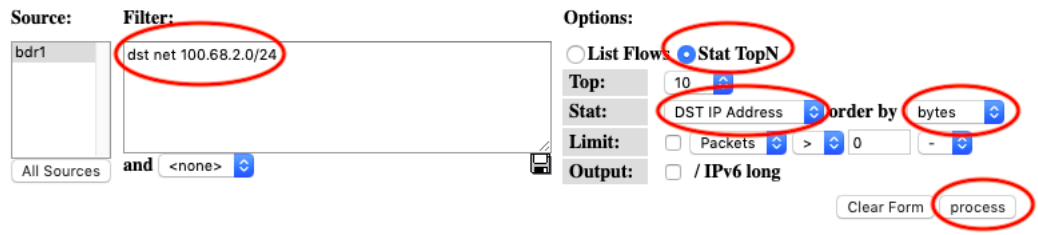
To do this, Stat "DST IP Address", order by "bytes".



Group flows by DST IP Address

This is now much closer to what we want: there is one line for each destination IP address, and they are ordered by total bytes, largest first.

But there is still one problem - can you see what it is? We are seeing a mixture of inbound flows (where the destination IP is inside our network) and outbound flows (where the destination IP is on the Internet). We are only interested in the inbound flows, so apply a filter which shows only traffic to your group's network: "dst net 100.68.X.0/24" (replacing X with your group number)



Flows to local network, grouped by DST IP Address

```
** nfdump -M /var/nfsen/profiles-data/live/bdr1  -T  -R 2019/11/14/nfcapd.201911141455:2019/11/14/nfcapd.201911141
nfdump filter:
dst net 100.68.2.0/24
Top 10 Dst IP Addr ordered by bytes:
Date first seen          Duration Proto      Dst IP Addr   Flows(%)      Packets(%)      Bytes(%)        pps
2019-11-14 14:56:45.123  3758.231 any        100.68.2.130    113( 6.6)     4925(45.7)   406694(45.3)        1
2019-11-14 14:57:01.694  3616.846 any          100.68.2.1    792(46.0)     2815(26.1)   229957(25.6)        0
2019-11-14 14:55:07.322  3738.637 any        100.68.2.131    269(15.6)      894( 8.3)    87250( 9.7)        0
2019-11-14 14:56:44.990  3720.825 any        100.68.2.132    217(12.6)      519( 4.8)    41524( 4.6)        0
2019-11-14 14:55:59.703  3773.848 any        100.68.2.133    191(11.1)      488( 4.5)    39168( 4.4)        0
2019-11-14 14:55:39.550  3790.661 any          100.68.2.2     62( 3.6)      317( 2.9)    26172( 2.9)        0
2019-11-14 14:56:09.943  3653.007 any        100.68.2.134     26( 1.5)      273( 2.5)    22620( 2.5)        0
2019-11-14 14:56:45.173  3618.361 any        100.68.2.136     26( 1.5)      273( 2.5)    22620( 2.5)        0
2019-11-14 14:56:28.955  3634.016 any        100.68.2.135     26( 1.5)      273( 2.5)    22620( 2.5)        0

Summary: total flows: 1722, total bytes: 898625, total packets: 10777, avg bps: 1864, avg pps: 2, avg bpp: 83
```

Output: Flows to local network, grouped by DST IP Address

At last we have what we want. The first record you see should tell you the local machine which has downloaded the most data in the period selected.

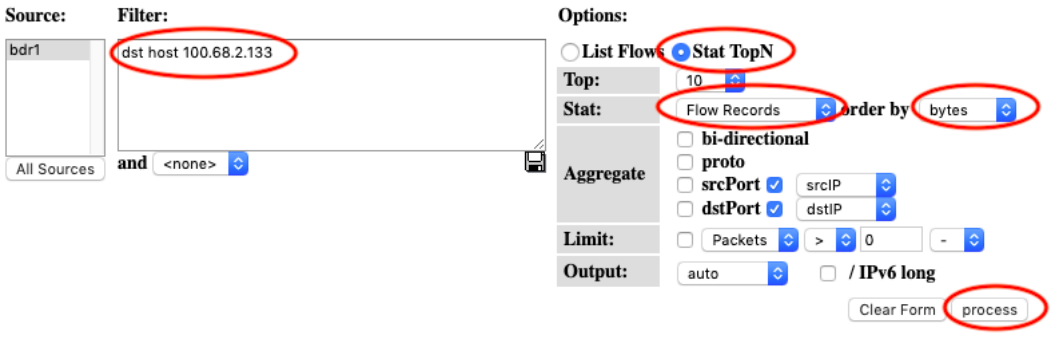
Outbound traffic grouped by sender IP address

Question: what changes would you have to make to this query to find out which machines in your network are *uploading* the most data to the Internet?

Analysing traffic to a single host

Now that we know which host has downloaded the most data, we might want to see where it has been downloading from.

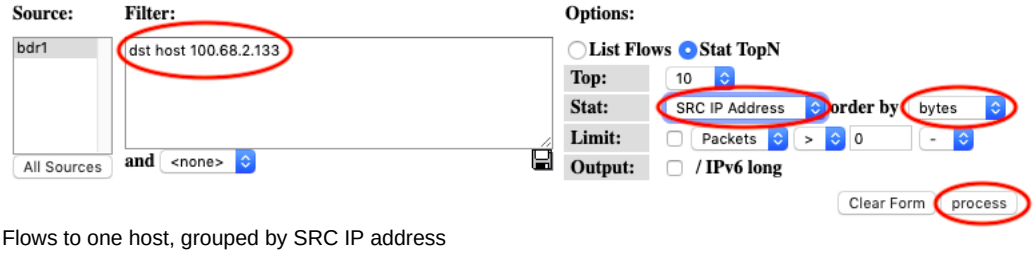
Let's start by looking at the top flows to that host. Change the filter to "dst host 100.68.X.Y" (the IP address you just found). Then select Stat "Flow Records", order by "bytes", and `process`.



Largest flows to one host

You should now see the flows inbound to that host, largest first. But again, we're only seeing large individual flows; a collection of small flows may add together to a large amount of traffic.

Since we are only looking at flow records to one particular destination IP address, we can group these records by source IP address.



Flows to one host, grouped by SRC IP address

```
** nfdump -M /var/nfsen/profiles-data/live/bdr1  -T  -R 2019/11/14/nfcapd.201911141455:2019/11/14/nfcapd.201911141
nfdump filter:
dst host 100.68.2.133
Top 10 Src IP Addr ordered by bytes:
Date first seen          Duration Proto      Src IP Addr   Flows(%)     Packets(%)       Bytes(%)         pps
2019-11-14 14:56:44.739  3618.353 any     192.168.122.250    13( 6.8)      260(53.3)     21840(55.8)        0
2019-11-14 15:15:54.761  2508.731 any       162.159.200.1    34(17.8)       52(10.7)      3952(10.1)        0
2019-11-14 15:15:54.561  2511.904 any     162.159.200.123    30(15.7)       47( 9.6)      3572( 9.1)        0
2019-11-14 15:01:59.270  3414.281 any      91.189.89.199     35(18.3)       40( 8.2)      3040( 7.8)        0
2019-11-14 15:17:18.225  2423.328 any      91.189.89.198     34(17.8)       35( 7.2)      2660( 6.8)        0
2019-11-14 15:19:26.192  2204.415 any       173.249.0.34     22(11.5)       28( 5.7)      2128( 5.4)        0
2019-11-14 14:55:59.703  1112.533 any      85.199.214.99     17( 8.9)       17( 3.5)      1292( 3.3)        0
2019-11-14 15:15:56.645    14.633 any       91.189.94.4       1( 0.5)        4( 0.8)       304( 0.8)        0
2019-11-14 15:03:36.063     0.000 any       212.18.3.19       1( 0.5)        1( 0.2)        76( 0.2)        0
2019-11-14 15:17:15.290     0.000 any      195.43.74.123      1( 0.5)        1( 0.2)        76( 0.2)        0

Summary: total flows: 191, total bytes: 39168, total packets: 488, avg bps: 83, avg pps: 0, avg bpp: 80
```

Output: Flows to one host, grouped by SRC IP address

And now we have one row for each IP address this host has been downloading from, with the total number of bytes downloaded from each IP, largest total first.

IP address information

By clicking on an IP address, you will get some information from reverse DNS and whois.

Whois information

Additional exercise: aggregating flows

NfSen offers some other ways to summarise the flows, using the Aggregate checkboxes. In this example we'll look again at traffic inbound to your network.

When you click one or more of the Aggregate boxes, NfSen combines all flows that share the same values of the attribute(s) you have selected.

To start this exercise, set the filter to "dst net 100.68.X.0/24" (X = your group). Select "Stat TopN", Stat "Flow Records", order by "bytes". Then try the following aggregates, remembering to click `process` after each one.

- Check "proto". You should get just one row each for TCP, UDP and ICMP, showing the total amount of traffic using each protocol. Sometimes this may show other protocols are active on your network (e.g. protocol 50 = IPSEC ESP; in Linux the file `/etc/protocols` has a list of them)
- Check both "proto" and "srcPort". This tells NfSen to combine together flows which have the same proto *and* the same srcPort. Depending on what activity has been going on, you may see one line giving the total for TCP port 80, one line for TCP port 443, one line for UDP port 53, and so on.
- Check "srcIP" by itself. This gives one row for each distinct source IP address, and is the same as selecting Stat SRC IP.
- Check both "srcIP" and "dstIP". You will get one row for each unique pair of srcIP and dstIP seen, with the total traffic between those two endpoints.

How would you change the filter to look at outbound traffic, rather than inbound traffic?

If you have a router with a full BGP table, you can aggregate netflow records by AS number. This is a useful way to find out what networks you are exchanging the most traffic with.