

Using NfSen to identify top talkers - Network Monitoring and Management

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 one of your PCs, and that PC is running NfSen to collect this data. If you are working in a pair, then you should both point your web browser to whichever PC is receiving the flows:

<http://db.campusX.ws.nsrc.org/nfsen/nfsen.php>

Generate some traffic

Firstly, we need to generate some traffic passing through your router. On either of your PCs (it doesn't have to be the one running NfSen), login and type the following commands:

```
$ cd /tmp
$ wget http://www.ws.nsrc.org/downloads/BigFile
$ rm BigFile
```

It will take around 5 minutes before this shows as a spike in NfSen.

Exploring flow records

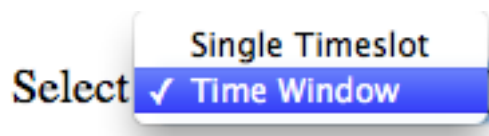
Now let's use NfSen to explore the traffic flows in the network, with the aim of finding out who was been downloading the most data. Look carefully at the output generated at each step - ask an instructor to explain if you don't understand what you see.

Navigate to Detail page

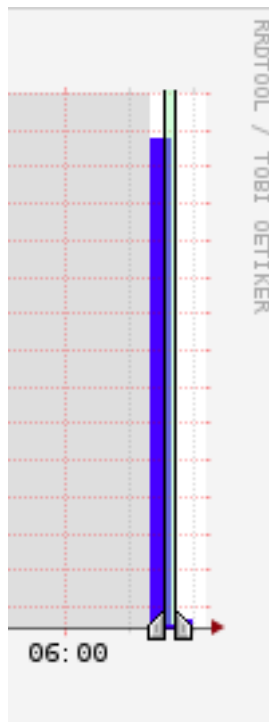
The NfSen home page shows a matrix of graphs: flows per second on the left, packets per second in the middle, bits per second on the right. Click on the top-right graph (bits per second, one day view) to get to the Detail page.

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:
✔ rtr1	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

Figure 1: 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.

Options:

☒ **List Flows** ☐ **Stat TopN**

Limit to: **Flows**

Aggregate

☐ **bi-directional**

☐ **proto**

☐ **srcPort**

☐ **dstPort**

Sort: ☐ **start time of flows**

Output: ☐ **/ IPv6 long**

Figure 2: 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:

However it’s still too much work to wade through this looking for interesting

Options:

☒ **List Flows** ☐ **Stat TopN**

Limit to: **Flows**

Aggregate

☒ **bi-directional**

☐ **proto**

☐ **srcPort**

☐ **dstPort**

Sort: ☐ **start time of flows**

Output: ☐ **/ IPv6 long**

Figure 3: Bi-directional flows

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 your campusX.srvY i.e 100.68.1.131)

The screenshot shows the NfSen interface. On the left, under 'Source:', there is a list with 'rtr1' selected. Below it is a button 'All Sources'. To the right of 'Source:' is the 'Filter:' section, which contains a text box with 'host 10.10.1.1' entered. Below the text box is a button 'and' followed by a dropdown menu showing '<none>'. To the right of the 'Filter:' section is the 'Options:' section. It has two radio buttons: 'List Flows' (selected) and 'Stat TopN'. Below these are several settings: 'Limit to:' with a value of '20' and a unit of 'Flows'; 'Aggregate' section with checkboxes for 'bi-directional' (unchecked), 'proto' (unchecked), 'srcPort' (unchecked), and 'dstPort' (unchecked); 'Sort:' with a checkbox for 'start time of flows' (unchecked); and 'Output:' with a dropdown set to 'auto' and a checkbox for '/ IPv6 long' (unchecked). At the bottom right of the 'Options:' section are two buttons: 'Clear Form' and 'process', with the 'process' button highlighted with a red circle.

Figure 4: 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**

The screenshot shows the NfSen interface. On the left, under 'Source:', there is a list with 'rtr1' selected. Below it is a button 'All Sources'. To the right of 'Source:' is the 'Filter:' section, which is empty. Below the text box is a button 'and' followed by a dropdown menu showing '<none>'. To the right of the 'Filter:' section is the 'Options:' section. It has two radio buttons: 'List Flows' (unchecked) and 'Stat TopN' (selected). Below these are several settings: 'Top:' with a value of '10'; 'Stat:' with a dropdown set to 'Flow Records'; 'Order by:' with a dropdown set to 'bytes'; 'Aggregate' section with checkboxes for 'bi-directional' (unchecked), 'proto' (unchecked), 'srcPort' (unchecked), and 'dstPort' (unchecked); 'Limit:' with a checkbox for 'Packets' (unchecked) and a value of '0'; and 'Output:' with a dropdown set to 'auto' and a checkbox for '/ IPv6 long' (unchecked). At the bottom right of the 'Options:' section are two buttons: 'Clear Form' and 'process', with the 'process' button highlighted with a red circle.

Figure 5: Find top flows by bytes

```

** nfdump -M /var/nfsen/profiles-data/live/rtr1 -T -R 2013/07/17/nfcapd.201307172050:2013/07/17/nfcapd.201307172
nfdump filter:
any
Verify map id 0: ERROR: Expected 7 elements in map, but found 2!
Aggregated flows 4194
Top 10 flows ordered by bytes:
Date first seen      Duration Proto      Src IP Addr:Port      Dst IP Addr:Port      Packets      Bytes      Flows
2013-07-17 18:34:59.964 8104.200 TCP        10.10.0.98:22 ->    10.10.0.241:56511     55346       31.8 M     1
2013-07-17 18:34:59.964 8104.200 TCP        10.10.0.241:56511 ->    10.10.0.98:22         29340       1.6 M      1
2013-07-17 20:28:46.766 1272.078 TCP        10.10.0.98:22 ->    10.10.0.241:56517     2669       389844     1
2013-07-17 20:28:46.766 1272.078 TCP        10.10.0.241:56517 ->    10.10.0.98:22         3383       224316     1
2013-07-17 20:50:29.950 15.832 TCP        10.10.0.98:80 ->    10.10.0.241:37764      57         73003      1

```

Figure 6: 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".

Options:

☐ List Flow
 ☒ **Stat TopN**

Top:

Stat: **order by**

Limit: ☐ Packets

Output: ☐ / IPv6 long

Figure 7: 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)

The screenshot shows the nfdump GUI with the following settings:

- Source:** rtr1
- Filter:** dst net 10.10.0.0/16
- Options:**
 - ☒ List Flows
 - ☒ Stat TopN
 - Top:** 10
 - Stat:** DST IP Address
 - order by:** bytes
 - Limit:** Packets > 0
 - Output:** /IPv6 long
- Buttons:** Clear Form, process

Figure 8: Flows to local network, grouped by DST IP Address

```
** nfdump -M /var/nfsen/profiles-data/live/gw-rtr -T -R 2013/04/17/nfcapd.201304170855:2013/04/17/nfcapd.201304171215 -n 10 -
nfdump filter:
dst net 10.10.0.0/16
Top 10 Dst IP Addr ordered by bytes:
Date first seen      Duration Proto  Dst IP Addr  Flows(%)  Packets(%)  Bytes(%)  pps  bps  bpp
2013-04-16 11:12:42.978 90437.613 any 10.10.0.135  92280(44.6)  1.9 M(41.8)  551.6 M(20.4)  20  48791  290
2013-04-17 06:55:42.339 19428.094 any 10.10.0.121  3924( 1.9)  303950( 6.7)  366.6 M(13.5)  15  150948  1206
2013-04-17 06:43:13.857 20201.599 any 10.10.0.115  2436( 1.2)  206384( 4.5)  288.9 M(10.7)  10  114424  1400
2013-04-17 08:52:41.704 12178.594 any 10.10.0.118  1044( 0.5)  111910( 2.5)  159.8 M( 5.9)  9  104992  1428
2013-04-16 10:56:01.483 91435.087 any 10.10.0.110  10446( 5.0)  192597( 4.2)  194.4 M( 5.7)  2  13512  801
```

Figure 9: 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**.

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.

Source: rtr1

Filter: dst host 10.10.1.1

Options:

☐ List Flow ☒ Stat TopN

Top: 10

Stat: Flow Records Order by: bytes

Aggregate:

☐ bi-directional

☐ proto

☐ srcPort srcIP

☐ dstPort dstIP

Limit: ☐ Packets > 0 -

Output: auto ☐ / IPv6 long

Clear Form process

Figure 10: Largest flows to one host

Source: rtr1

Filter: dst host 10.10.1.1

Options:

☐ List Flow ☒ Stat TopN

Top: 10

Stat: SRC IP Address Order by: bytes

Limit: ☐ Packets > 0 -

Output: ☐ / IPv6 long

Clear Form process

Figure 11: Flows to one host, grouped by SRC IP address

```
** nfdump -M /var/nfsen/profiles-data/live/gw-rtr -T -R 2013/04/17/nfcapd.201304170855:2013/04/17/nfcapd.201304171215 -n 10 -
nfdump filter:
dst host 10.10.0.135
Top 10 Src IP Addr ordered by bytes:
Date first seen Duration Proto Src IP Addr Flows(%) Packets(%) Bytes(%) pps bps bpp
2013-04-17 09:59:37.965 7177.308 any 86.135.63.204 70( 0.1) 133384( 7.0) 166.0 M(30.1) 18 185002 1244
2013-04-17 11:58:22.389 652.388 any 155.232.240.14 16( 0.0) 41268( 2.2) 57.4 M(10.4) 63 703269 1389
2013-04-17 09:49:27.947 4725.000 any 39.52.237.91 4( 0.0) 38278( 2.0) 46.9 M( 8.5) 8 79376 1224
2013-04-17 10:02:47.530 2510.000 any 109.65.3.106 4( 0.0) 35506( 1.9) 36.9 M( 6.7) 14 117603 1039
2013-04-17 11:00:02.692 4285.997 any 168.122.196.248 8( 0.0) 21956( 1.2) 22.9 M( 6.0) 5 61352 1497
```

Figure 12: 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.

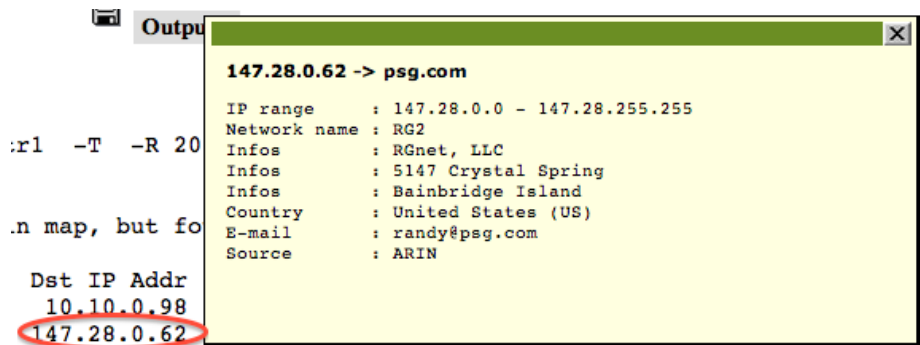


Figure 13: Whois information

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.