Should network operators hop on the data plane?

NANOG 83

Max Resing
contact@maxresing.de
University of Twente

Who am I?

Max Resing

- MSc Computer Science
  Specialization:
  Data Science
- Part-time DevOps
  Engineer

Background

Problem: Internet scanners

Solution: IP blocklists

Hypothesis

Global diversity is insufficient to detect all scanners if the sensor infrastructure operates purely in cloud environments

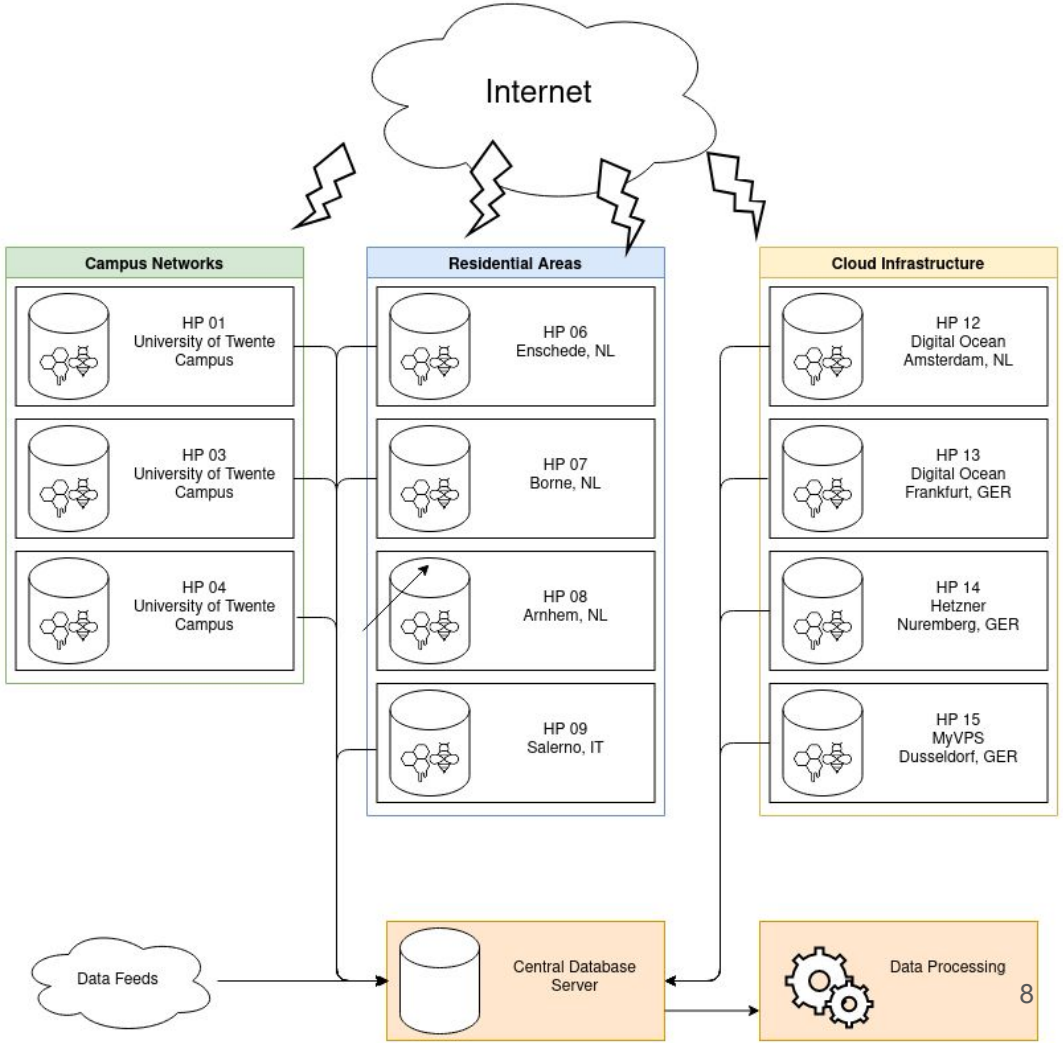
Goal

- Track scanners of different kinds of networks
- Compare origins of scanning activity
- Is the type of network relevant to identify scanners?

Setup & Infrastructure

Infrastructure

- 3 network types
- 11 honeypots



Honeypot

- Debian-based
- OpenSSH, Telnet, VNC

- Logging Period

May 2021						
					01	02
03	04	05	06	07	08	09
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

Results

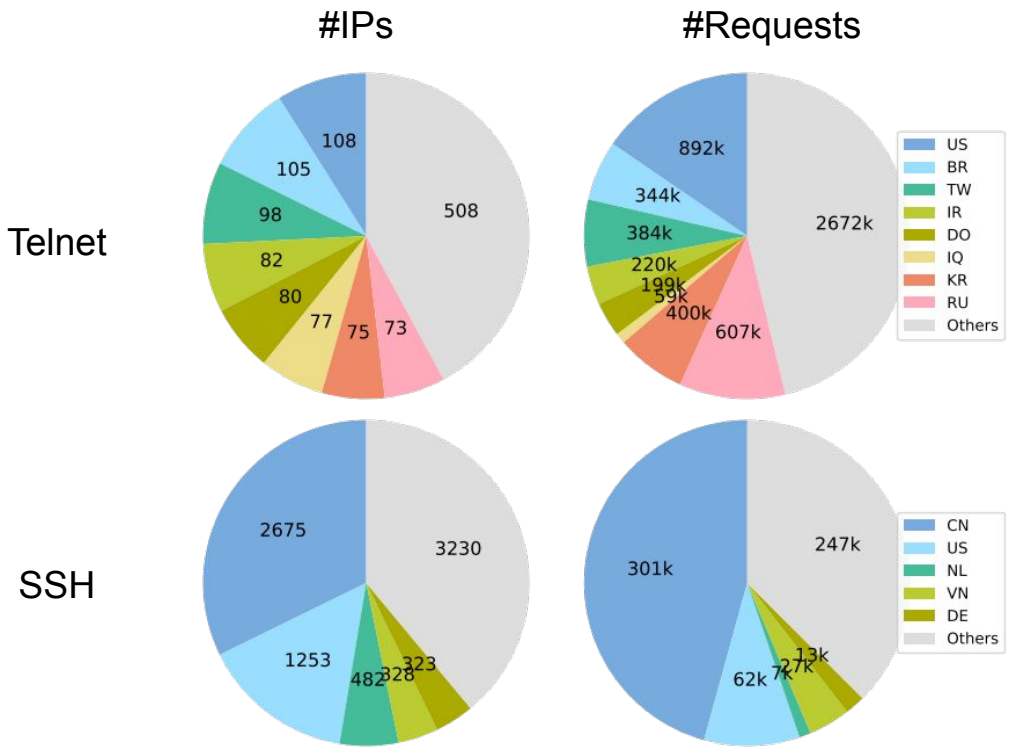
6,630,498 requests

7,182 IPs

What did we do with it?

- Geographical activity
- Importance of TOR traffic
- Temporal activity
- Temporal activity per timezone
- Coverage of Dataplane feeds

Geographical Activity of Scanners



Geographical Activity of Scanners

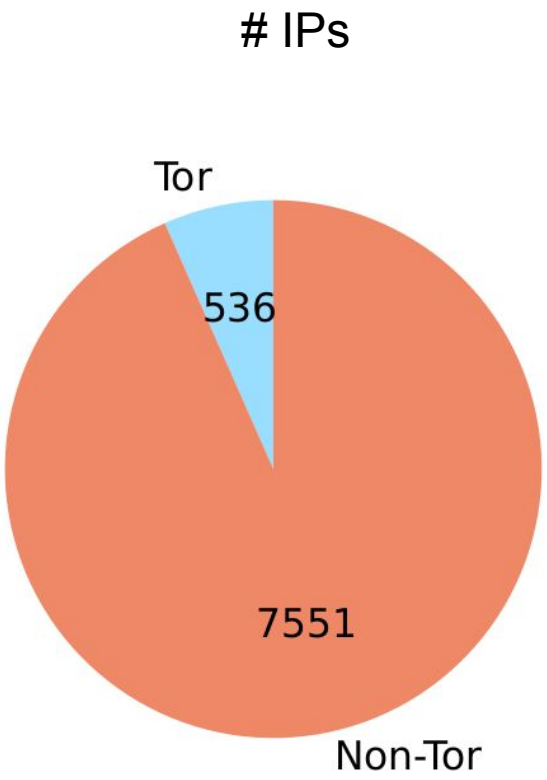
Telnet

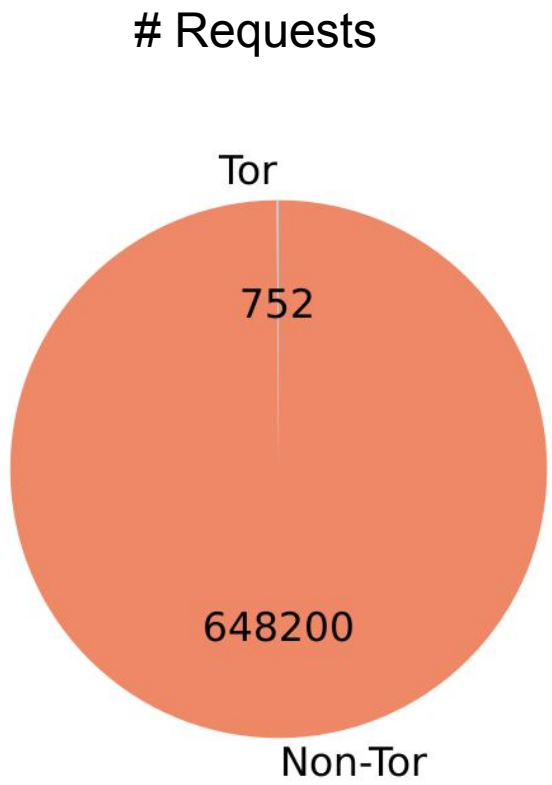
- distributed equally over the globe

SSH

- China → University campus & cloud networks
- US → Much more interest in residential areas
- Netherlands → Many IPs, few requests

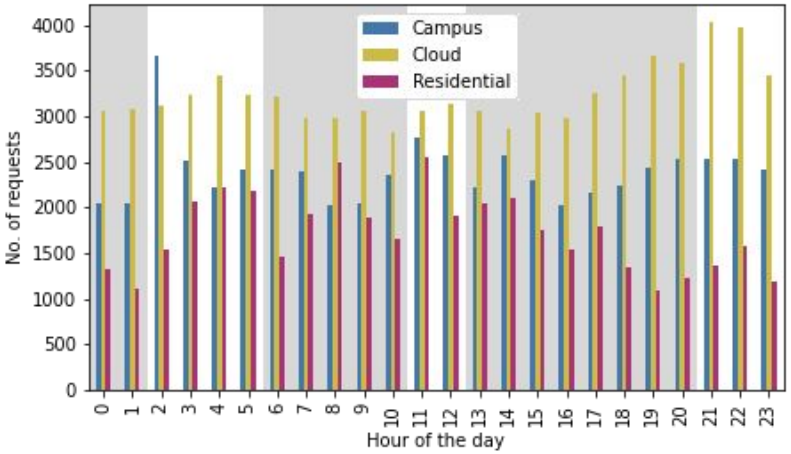
Importance of TOR Traffic





Temporal Analysis

- No weekly pattern
- Slight peaks during certain hours



Temporal Activity per Timezone

SSH

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
-07:00	1404	824	978	1161	1790	1489	875	968	1317	924	818	1145	1688	1419	1116	1400	1258	971	1159	1869	1051	1529	1015	1069
+03:00	2290	1936	1345	516	523	677	1158	937	337	457	541	1687	1134	1255	1540	1674	1228	868	1261	1423	1597	1565	1403	1458
+08:00	12446	13666	13911	14123	14910	15156	15138	15231	14260	13666	18029	14981	15005	14094	13460	13632	13017	13608	13123	13936	13882	14327	13298	12710

Telnet

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
-04:00	36563	39332	39565	39235	40154	42464	44745	45686	47895	48587	46700	46230	43652	43887	45535	44660	42413	40813	41645	42114	38433	37367	36146	34998
+00:00	0	5	0	0	906	986	987	987	980	948	868	860	878	257	0	0	0	0	0	0	9	76	0	0
+02:00	48322	47841	47619	47757	46916	49911	51366	52318	55209	54622	52668	52225	54026	58012	54991	51775	49826	49460	52480	52160	49158	47881	47951	45706

*Timezones in GMT

Coverage
Dataplane Feeds

Coverage of SSH scanners in Feeds

SSH	Total	Covered (%)
Cloud	3,667	90.8
Cloud only		
Campus	3,695	97.0
Campus only		
Residential	875	52.0
Residential only		

Coverage of SSH Scanners in Feeds

SSH		Total	Covered (%)
Cloud		3,667	90.8
Cloud only	~ 79%	2,895	88.3
Campus		3,695	97.0
Campus only	~ 46%	1,689	95.8
Residential		875	52.0
Residential only	~ 67%	590	29.0

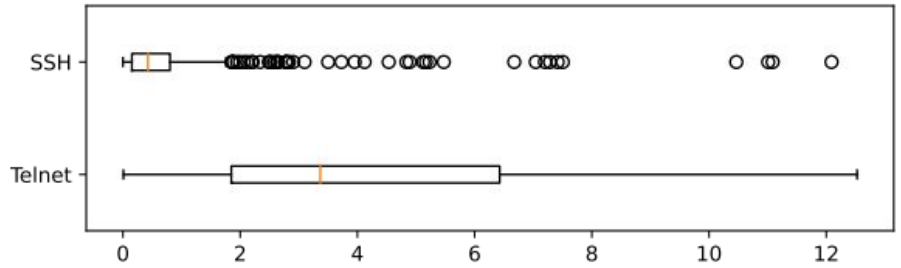
Coverage of Telnet Scanners in Feeds

Telnet	Total	Covered (%)
Cloud	416	15.4
Cloud only		
Campus	371	7.5
Campus only		
Residential	431	10.9
Residential only		

Coverage of Telnet Scanners in Feeds

Telnet		Total	Covered (%)
Cloud		416	15.4
Cloud only	~ 97%	402	14.9
Campus		371	7.5
Campus only	~ 97%	360	6.9
Residential		431	10.9
Residential only	~ 97%	418	10.9

Delayed Discovery of Scanners



*Days passed until IP showed up in feeds

Updates

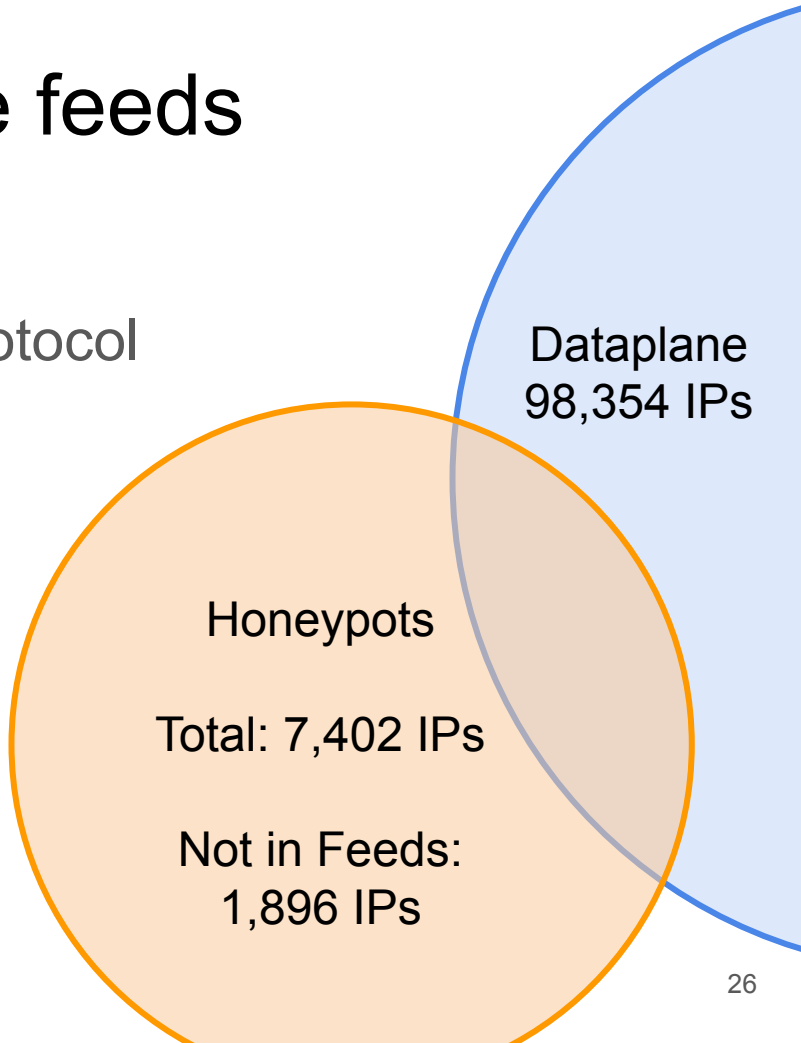
Skewed IPs

- Bogus IP addresses
- ~ 3% requests

- **telnetd** logged partially rDNS, partially IPs
- Parsing raw log data parsed some rDNS entries
- Affects ~ 5.7% of telnet scanner IPs

Comparison with Dataplane feeds

- Extended period (4 weeks)
- No consideration of network type/protocol



Conclusion

- Scanners target certain types of networks
- SSH & Telnet scanners are of different nature

Most importantly:

- Blocklist providers require qualitative diversity to discover all scanners

Thank you!

Special thanks to
John Kristoff

Max Resing

contact@maxresing.de

www.maxresing.de