

Real-world DSCP Traversal Measurements

(Doomsday edition, by popular request)

Ana Custura

ana@erg.abdn.ac.uk

draft-ietf-tsvwg-dscp-considerations-05

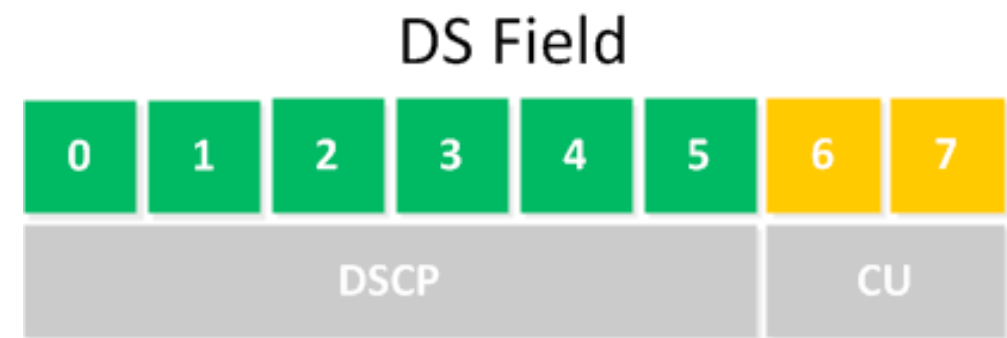
Background

- IP packets had a ToS Byte

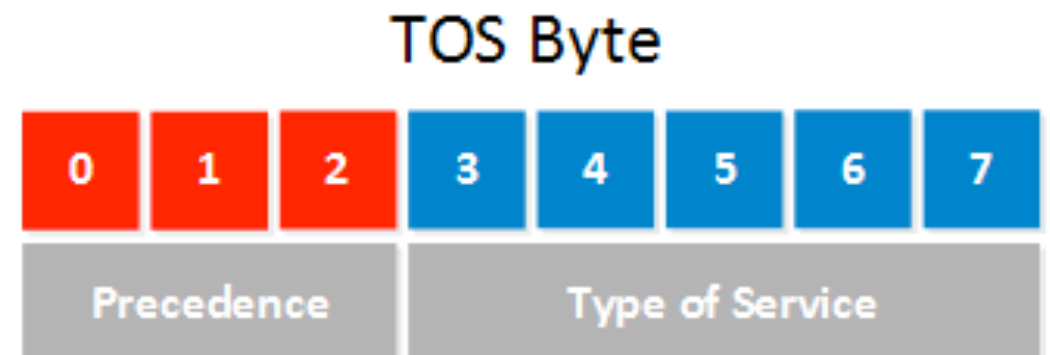


- DSCP = Differentiated Services Code Point
 - IP nodes on path associate DSCPs with a Per-Hop Behaviour
 - Helps classify traffic and provide QoS

- DS field/IPv6 TC now:



- ToS Byte pre-1998, RFC781:



Background

- DiffServ routers operate on a 6-bit field, NOT on the former 3-bit Precedence field - RFC 2474

- Or at least, in theory :)

- RFC 2474 keeps compatibility by defining “CS” DSCPs:

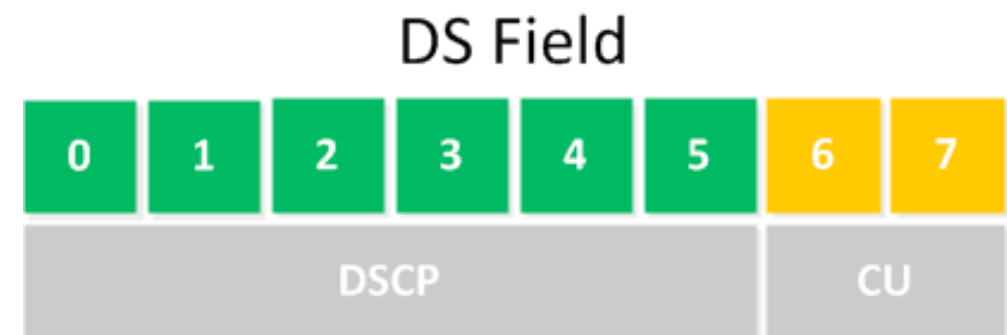
000 -> 000 000 (CS0) 0

001 -> 001 000 (CS1) 8

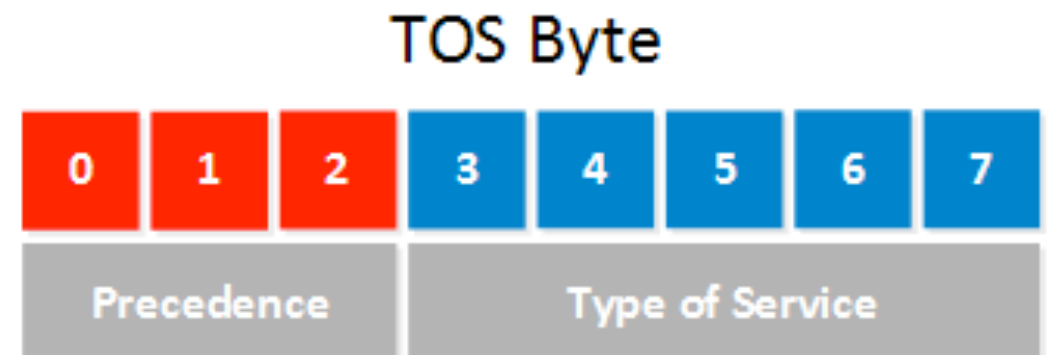
.....

111 -> 111 000 (CS7) 56

- DS field/IPv6 TC now:



- ToS Byte pre-1998, RFC781:



DSCP took off:

IANA Registrations & Reserved DSCPs

- ToS Byte Precedence: CS DSCPs in RFC 2474
- AF (Assured Forwarding) DSCPs in RFC 2597
- EF (Expedited Forwarding), RFC 3246
- Voice Admit, RFC 5865
- LE (Lower Effort), RFC 8622
- Experimental use: xxxx11

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

DSCP took off:

IANA Registrations & Reserved DSCPs

- ToS Byte Precedence: CS DSCPs in RFC 2474
- AF (Assured Forwarding) DSCPs in RFC 2597
- EF (Expedited Forwarding), RFC 3246
- Voice Admit, RFC 5865
- LE (Lower Effort), RFC 8622
- Experimental use: xxxx11

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

DSCP took off:

IANA Registrations & Reserved DSCPs

- ToS Byte Precedence: CS DSCPs in RFC 2474
- AF (Assured Forwarding) DSCPs in RFC 2597
- EF (Expedited Forwarding), RFC 3246
- Voice Admit, RFC 5865
- LE (Lower Effort), RFC 8622
- Experimental use: xxxx11

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

DSCP took off:

IANA Registrations & Reserved DSCPs

- ToS Byte Precedence: CS DSCPs in RFC 2474
- AF (Assured Forwarding) DSCPs in RFC 2597
- EF (Expedited Forwarding), RFC 3246
- Voice Admit, RFC 5865
- LE (Lower Effort), RFC 8622
- Experimental use: xxxx11

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

DSCP took off:

IANA Registrations & Reserved DSCPs

- ToS Byte Precedence: CS DSCPs in RFC 2474
- AF (Assured Forwarding) DSCPs in RFC 2597
- EF (Expedited Forwarding), RFC 3246
- Voice Admit, RFC 5865
- LE (Lower Effort), RFC 8622
- Experimental use: xxxx11

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

Are They Used?

- AF11, AF21, CS3, EF used by webserver & some DNS servers
- AF11,12,13 observed in mobile networks
- ICMP traffic marked with DSCP CS6
- DNS likes CS1, CS4, AF11
- Measured looking at server replies, and in passive network traces (see Appendix)

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

Are They Used?

- AF11, AF21, CS3, EF used by webserver & some DNS servers
- AF11,12,13 observed in mobile networks
- ICMP traffic marked with DSCP CS6
- DNS likes CS1, CS4, AF11
- Measured looking at server replies, and in passive network traces (see Appendix)

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

Are They Used?

- AF11, AF21, CS3, EF used by webserver & some DNS servers
- AF11,12,13 observed in mobile networks
- ICMP traffic marked with DSCP CS6
- DNS likes CS1, CS4, AF11
- Measured looking at server replies, and in passive network traces (see Appendix)

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

Are They Used?

- AF11, AF21, CS3, EF used by web servers & some DNS servers
- AF11,12,13 observed in mobile networks
- ICMP traffic marked with DSCP CS6
- DNS likes CS1, CS4, AF11
- Measured looking at server replies, and in passive network traces (see Appendix)

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

Are They Used?

- AF11, AF21, CS3, EF used by web servers & some DNS servers
- AF11,12,13 observed in mobile networks
- ICMP traffic marked with DSCP CS6
- DNS likes CS1, CS4, AF11
- Measured looking at server replies, and in passive network traces (see Appendix)

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (i)

- could they be used end-to-end?

- Internet core study shows some routers still use former IP precedence bits (2015-2017):
 - Clearing only 3 MSB of DSCPs:
ToS Precedence Bleaching
- Supported by packet trace analysis (2018, 2019)
- Validated in edge networks (2021)
- **..all DSCPs < 8 have higher end-2-end traversal**

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (i)

- could they be used end-to-end?

- Internet core study shows some routers still use former IP precedence bits (2015-2017):
 - Clearing only 3 MSB of DSCPs:
ToS Precedence Bleaching
- Supported by packet trace analysis (2018, 2019)
- Validated in edge networks (2021)
- **..all DSCPs < 8 have higher end-2-end traversal**

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (ii)

- could they be used end-to-end?

- AFx1 ToS Precedence Bleach to DSCP 2 - currently unusable - a lot of traffic aggregated to this
- DSCP 4: SSH clients set this code point for SSH traffic :-)
- DSCP 6: same issue as DSCP 2
- DSCP 57,49,41,33,25,17,9 aggregate to LE (DSCP 1)
- DSCPs 3 and 7 (exp pool) left, and DSCP 5

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (ii)

- could they be used end-to-end?

- AFx1 ToS Precedence Bleach to DSCP 2 - currently unusable - a lot of traffic aggregated to this
- DSCP 4: SSH clients set this code point for SSH traffic :-)
- DSCP 6: same issue as DSCP 2
- DSCP 57,49,41,33,25,17,9 aggregate to LE (DSCP 1)
- DSCPs 3 and 7 (exp pool) left, and DSCP 5

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (ii)

- could they be used end-to-end?

- AFx1 ToS Precedence Bleach to DSCP 2 - currently unusable - a lot of traffic aggregated to this
- DSCP 4: SSH clients set this code point for SSH traffic :-)
- DSCP 6: same issue as DSCP 2
- DSCP 57,49,41,33,25,17,9 aggregate to LE (DSCP 1)
- DSCPs 3 and 7 (exp pool) left, and DSCP 5

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

The Problem (ii)

- could they be used end-to-end?

- AFx1 ToS Precedence Bleach to DSCP 2 - currently unusable - a lot of traffic aggregated to this
- DSCP 4: SSH clients set this code point for SSH traffic :-)
- DSCP 6: same issue as DSCP 2
- DSCP 57,49,41,33,25,17,9 aggregate to LE (DSCP 1)
- DSCPs 3 and 7 (exp pool) left, and DSCP 5

56 CS7	57	58	59	60	61	62	63
48 CS6	49	50	51	52	53	54	55
40 CS5	41	42	43	44 VA	45	46 EF	47
32 CS4	33	34 AF41	35	36 AF42	37	38 AF43	39
24 CS3	25	26 AF31	27	28 AF32	29	30 AF33	31
16 CS2	17	18 AF21	19	20 AF22	21	22 AF23	23
8 CS1	9	10 AF11	11	12 AF12	13	14 AF13	15
0 CS0	1 LE	2	3	4	5	6	7

DSCP Traversal: 2015-2017

- Mobile Edge (Traceroutes for well-known DSCPs from 12 European mobile carriers, 2015-2016):
 - **ToS Precedence Bleaching** emerges as pathology
 - Most mobile networks remark to a single DSCP
 - AF 11, 12, 13 are popular choices
- Core (Traceroutes for all 64 DSCPs to 500 web servers, from 8 vantage points, 2017):
 - 80% end2end traversal rate for DSCPs<7
 - 55% end2end traversal rate for DSCPs >7
 - **ToS Precedence Bleaching** on up to 20% of paths

CAIDA Passive Traces

- Traffic recorded during 1 hour in January at Equinix DC in NY
- OC192 backbone link (9953 Mbps) of a Tier 1 ISP between NY and Sao Paulo
- Traffic is bidirectional and anonymised preserving prefixes
- We looked at DSCP markings

	2018	2019
IPv4 Packets	9,568,663,465	7,937,877,712
IPv6 Packets	316,007,533	234,393,206
TCP/UDP split	79 / 19 (2% other traffic)	73 / 24 (3% other traffic)

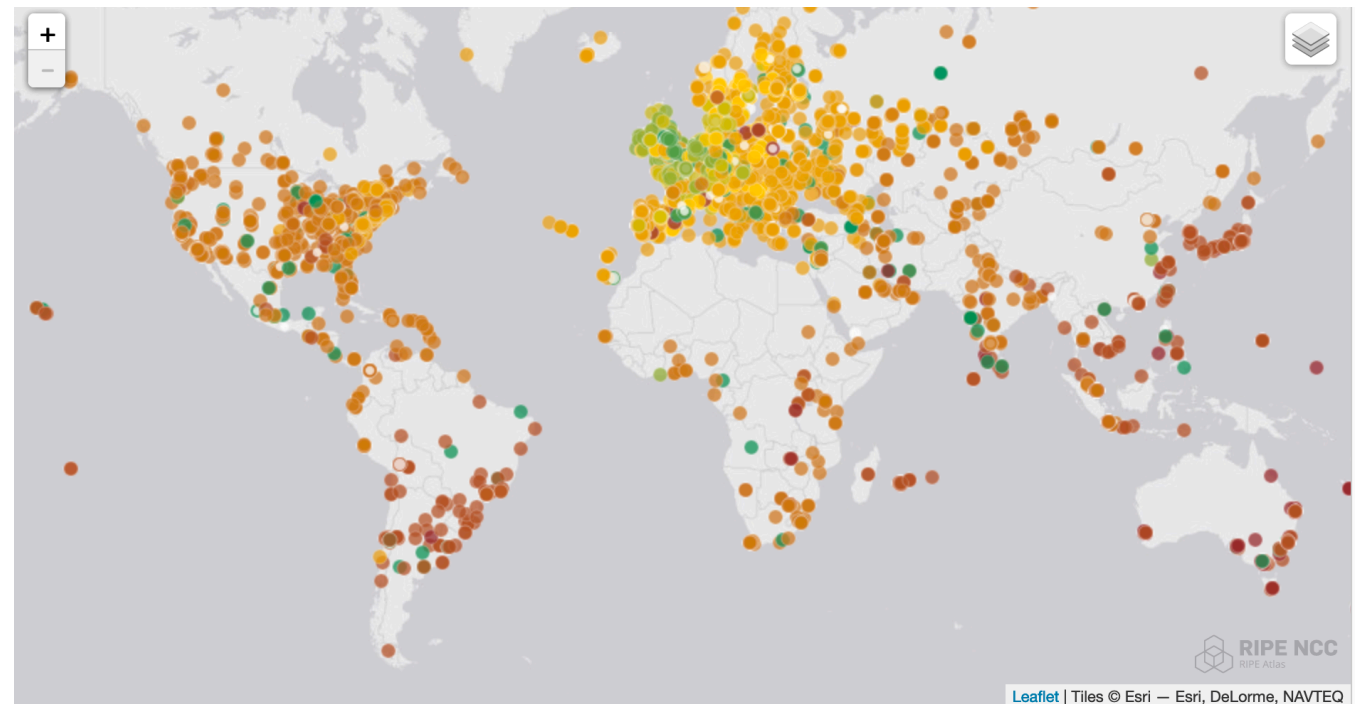
CAIDA Passive Traces

	2018 - Dir A	2018 - Dir B	2019 - Dir A	2019 - Dir B
All packets Per direction and year	BE: 91% 42, 44, 46: < 0.01% 2: 7% 6: 0.5% 4: 0.3% 5: 0.1%	BE: 79% 42: < 0.01% 2: 15% 5: 3.8% 4: 0.4% 6: 0.1%	BE: 78% 42: < 0.01% 2: 19% 4: 0.2% 5: 0.2% 6: 0.3%	BE: 81% 42: < 0.01% 2: 13% 5: 4.9% 4: 0.3% 6: 0.1%
	2018 - IPv4	2018 - IPv6	2019 - IPv4	2019 - IPv6
All packets Per protocol	BE: 82% 42, 44, 46: <=0.01% 2: 13% 5: 3% 4: 0.4% 6: 0.2%	BE: 90% 42, 44, 46: <0.01% 2: 7% 5: 0.03% 4: < 0.01% 6: < 0.01%	BE: 79% 42, 44, 46: <0.01% 2: 14% 5: 3.6% 4: 0.2% 6: 0.2%	BE: 94% 42, 44, 46: <0.01% 2: 4% 4: < 0.01% 5: < 0.01% 6: < 0.01%

- Traffic w/**DSCP 2** accounts for ~4-19%, present across direction, year and IP version split.
- Other **DSCPs (4, 5, 6)** seen less, with the exception of DSCP 5 in Dir B
- **DSCP 2** results from ToS bleaching of **AF11, AF21, AF31 and AF41**

RIPE Atlas traceroutes

- 10,000 RIPE Atlas probes, many in edge networks
- Sender DSCP set to 0, 42, 45, 2, 5,
- TCP traceroute to port 8080 at UoA server, June 2021



- DSCP 45 has a 20% traversal rate; DSCP 5 has a 25% traversal rate
- Bleaching to DSCP 0 happens in the first AS on up to 40% of paths
- **ToS bleaching still happening on up to 10% of paths**

What do we do with this data?

- Do we accept only 8 DSCPs can traverse the Internet e2e?
- **Considerations for Assigning new DSCPs - TSVWG**
- Helped guide discussion of allocation of a new DSCP for Non-Queue-Building Traffic:
 - Originally proposed 2 DSCPs (45 and 5)
 - Dual assignments (45/5) risk making DSCPs in the same column unusable
 - Only one DSCP (45) allocated in the end.

Operator Survey Time

- Can we make recommendations for future assignments?

https://a.custura.eu/pub/MNM2017_DSCPmobile.pdf

https://a.custura.eu/pub/ACM_CC2018_ExploringDSCP.pdf

<https://datatracker.ietf.org/doc/draft-ietf-tsvwg-dscp-considerations/>

- Questions?

<https://forms.microsoft.com/r/fAD5ckcRaw>



Appendix

DSCPs set by applications

(CAIDA dataset 2018)

- ICMP - DSCP 48 - 19.1%
- SSH - DSCP 4 - 31.5%;
- DNS - DSCP 5 - 8.1%;
- SIP - DSCP 5, 14%

Honorable mentions:

- TCP, ports 8001 and 8880 - DSCP 1 - 65% and 45% respectively
- TCP, port 9050 (Tor SOCKS port) - DSCP 2 - 49.9%;
- TCP and UDP, port 8999: **DSCP 5 - 16%**; (what is this??)

Server Reply Datasets - PATHSpider

DSCPs set by DNS servers/server edge

- DNS request to authoritative NSes for Alexa top 1M domains
- Replies from 15% servers with non-0 DSCP (v4 and v6)
- Popular code points are 8 (CS1), 32 (CS4), 10 (AF11), 14 (AF13), 6 (ToS Bleach of EF)

DSCP	Percentage
CS1	71%
CS4	6%
AF11	4.9%
AF13	3.22%
ToS Bleach EF	3%
1	1.9%
AF31	1.8%
AF21	1.7%

Server Reply Datasets - PATHSpider

DSCPs set by HTTP servers/server edge

- HTTP request to Alexa top 1M domains (over 400k IP addresses)
- Replies from 52762 (12.7%) servers with non-0 DSCP (v4)
- Popular code points are AF11, AF21, AF31 and CS3
- Of these, 1222 servers reply with DSCP 5

DSCP	Percentage
AF11	63.3
AF21	18.2
CS3	4.45
6 (ToS bl of EF)	2.45
5	2.31
AF31	1.96
1	1.63
CS6	1.47