IPv4 Address Sharing: Problem, Solutions, and Test results

draft-abdo-hostid-tcpopt-implementation
Authors: E. Abdo, J. Queiroz, M. Boucadair
draft-wing-nat-reveal-option
Authors: A. Yourtchenko, D. Wing

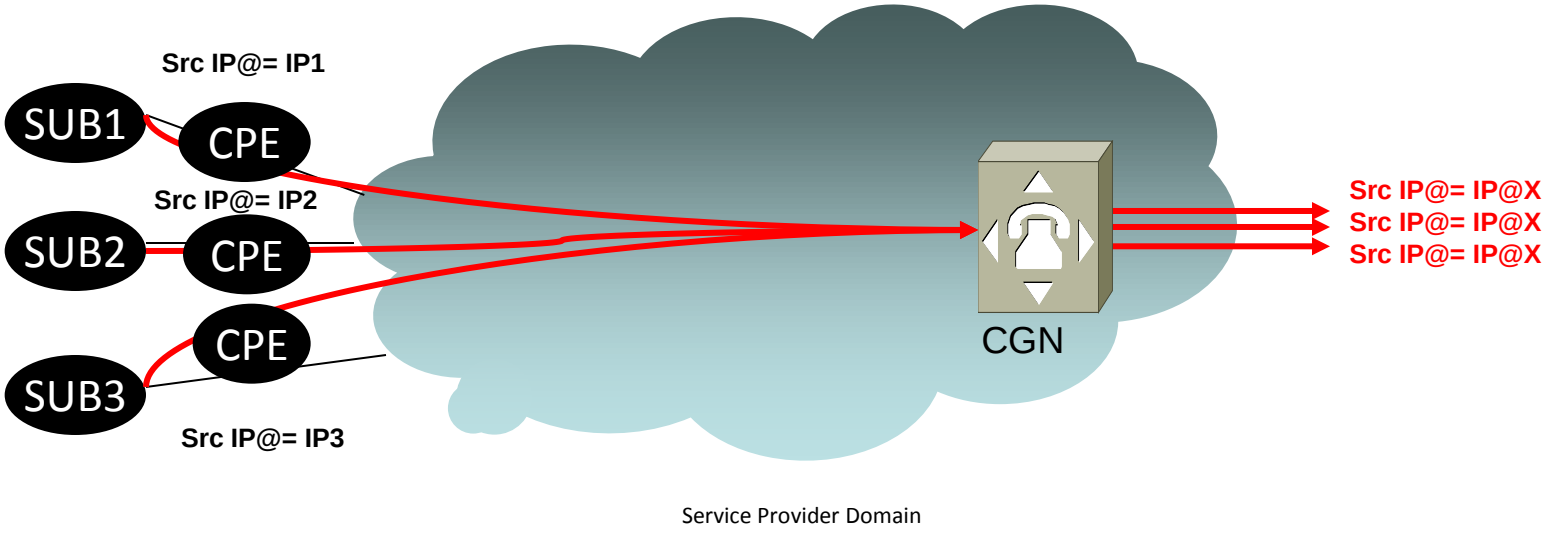
TCPM WG
IETF 84-Vancouver, July 2012

Presenter: J. Queiroz

Problem to be Solved

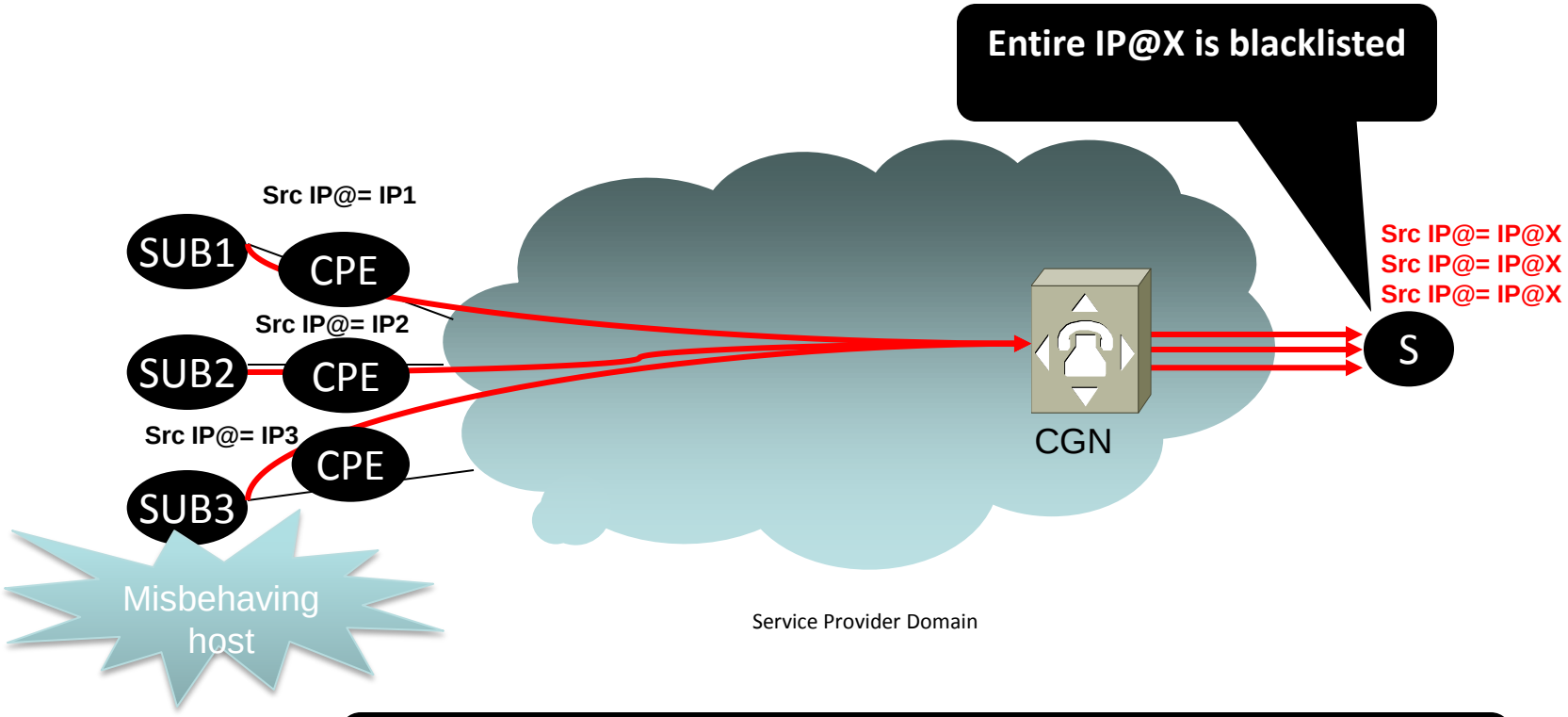
- **Context**
  - Public IPv4 address depletion
  - IPv4 service continuity should be maintained
    - Necessity of large scale address sharing
- **IPv4 address sharing solutions**
  - CGN/NAT64/DS-Lite/A+P/4rd/DIVI
  - Application proxies (e.g., HTTP proxies)
- **Issues with IPv4 address sharing**
  - Documented in RFC 6269
  - Issues for end-users, service providers, content providers and legal authorities
    - Specific use case that causes denial of service

Address Sharing



The internal and the external IP addresses may be of distinct address families (e.g., IPv4, IPv6): NAT44 or NAT64

Implicit Identification



Blacklisting a misbehaving user:
The server relies on the source IP address

All subscribers using the same address will be impacted:
Unhappy customers, calls to the hotline for the IP Network Provider ($$/mn, OPEX loss for the ISP)

Results from intarea-nat-reveal-analysis

solution tested in abdo-hostid-tcpopt-implementation

	UDP	TCP	HTTP	Encrypted traffic	Success Ratio	Possible performance impact	Modify OS TCP/IP stack is needed (*)	Deployable	Notes
IP Option	Yes	Yes	Yes	Yes	30%	High	Yes	Yes	
TCP Option	No	Yes	Yes	Yes	99%	Med to High	Yes	Yes	
IP-ID	Yes	Yes	Yes	Yes	100%	Low to Med	Yes	Yes	1
HTTP Header (XFF)	No	No	Yes	No	100%	Med to High	No	Yes	2
Proxy Protocol	No	Yes	Yes	Yes	Low	High	No	No	
Port Set	Yes	Yes	Yes	Yes	100%	NA	No	Yes	1,3
HIP					Low	NA	--	No	4,5

(1) Requires mechanism to advertise NAT is participating in this scheme (e.g., DNS PTR record)
(2) This solution is widely deployed
(3) When the port set is not advertised, the solution is less efficient.
(4) Requires the client and the server to be HIP-compliant and HIP infrastructure to be deployed
(5) If the client and the server are HIP-enabled, the address sharing function does not need to insert a user-hint. If the client is not HIP-enabled, designing the device that performs address sharing to act as a UDP/TCP-HIP relay is not viable.

(*) Server side

IP option, IP ID and Proxy Protocol are **broken**

HIP is not “widely” **deployed**

Port Set requires **coordination**

XFF is **largely deployed** in operational networks but still the address sharing function **needs to parse all applications messages**

TCP Option is superior to XFF since it is not specific to HTTP but what about **UDP**? Update the Servers OS **TCP/IP is required**

HOST_ID

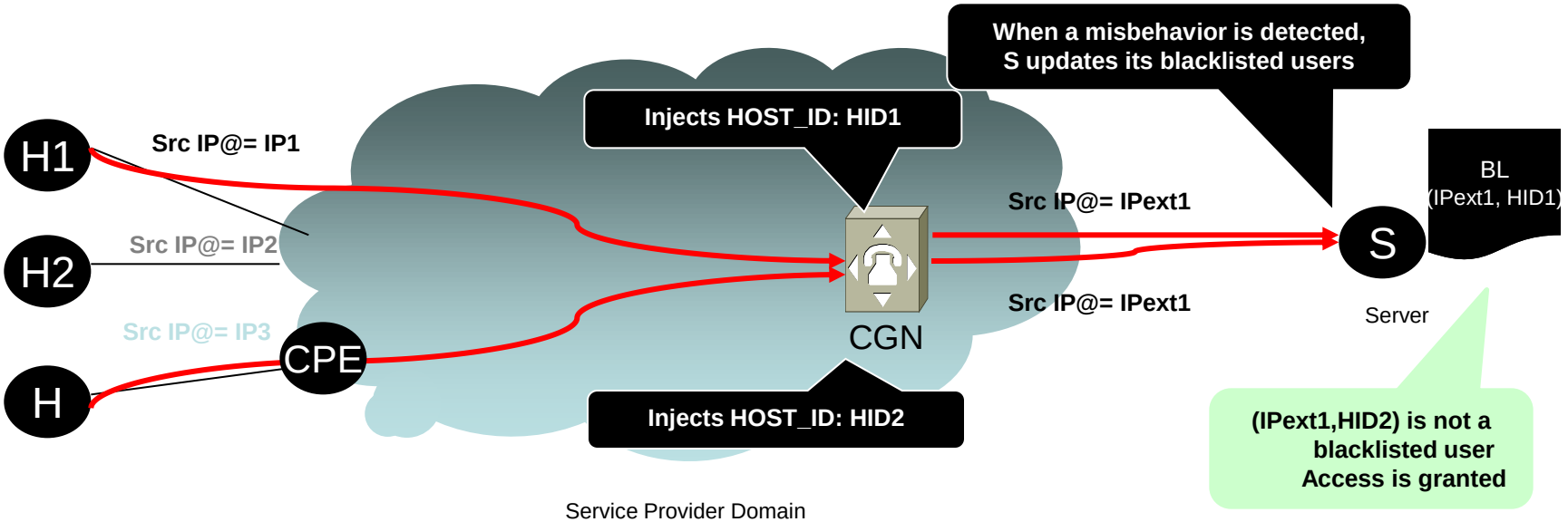
- **What is the HOST_ID?**
  - It must be **unique** to each user who shares the same global IPv4 address
  - Adding a HOST_ID does not “break” the privacy of the user
  - E.g. first bits of an IPv6 address, private IPv4 address, etc.
- **Who puts the HOST_ID?**
  - The address sharing function injects the HOST_ID when NAT operation is in process
    - The CPE can put the identification in the packet and the CGN checks it instead of injecting he information itself. The performance impact would be distributed between CPE and CGN
- **Where is the HOST_ID?**
  - If the HOST_ID is put at the IP level, all packets will have to bear the identifier
  - If it is put at a higher connection-oriented level, the identifier is only needed once in the session establishment phase
    - **E.g., TCP option**

HOST_ID as a TCP OPTION

- Original idea is documented in I-D.wing-nat-reveal-option
  - 4 bytes long
  - Denoted as HOST_ID_WING

- An additional TCP option format to convey a HOST_ID is also considered
  - 10 bytes long
  - Denoted as HOST_ID_BOUCADAIR
  - **Motivation**: cover also the load-balancer use case and provide richer functionality as Forwarded-For HTTP header

Illustrating Encountered Issues (Revisited)



Blacklisting a misbehaving user:
The server relies on the source IP address **& HOST_ID**

The server needs to be updated to:
(1) be able to extract the HOST_ID, (2) Enforce policies based on the HOST_ID, (3) log the HOST_ID

I-D.abdo-hostid-tcpopt-implementation

- Various combinations of the HOST_ID as TCP option were tested
  - HOST_ID_WING
    - HOST_ID_WING was also adapted to include 32 bits and 64 bits values
    - No particular impact on session establishment was observed
  - HOST_ID_BOUCADAIR (source port)
  - HOST_ID_BOUCADAIR (IPv4 address)
  - HOST_ID_BOUCADAIR (source port:IPv4 address)
  - HOST_ID_BOUCADAIR (IPv6 Prefix)

Main Tests' Objectives

1. Assess the validity of the HOST_ID TCP option approach
2. Assess the behavior of legacy TCP servers when receiving a HOST_ID TCP option
3. Assess the impact of injecting a HOST_ID TCP option on the time it takes to establish a connection
4. Assess the performance impact on the CGN device that has been configured to inject the HOST_ID TCP option

- All tests' results can be found in detail:
  I-D.abdo-hostid-tcpopt-implementation

Conclusions

- HOST_ID implementation is feasible and not complex
- No impact for HOST_ID options on TCP session establishment delay
- HTTP sessions success ratio is not significantly impacted by the presence of HOST_ID options (0.105% failures - WING)
- FTP session success ratio is slightly impacted by the presence of HOST_ID options (0.44% Connection failures)
- No impact for HOST_ID options on ISC-CGN performance
- Policies based upon HOST_ID contents were applied and tested successfully (log, deny, match, strip)
- Similar implementations on going (one regards open-source proxy software applications; and other under content provider environment)

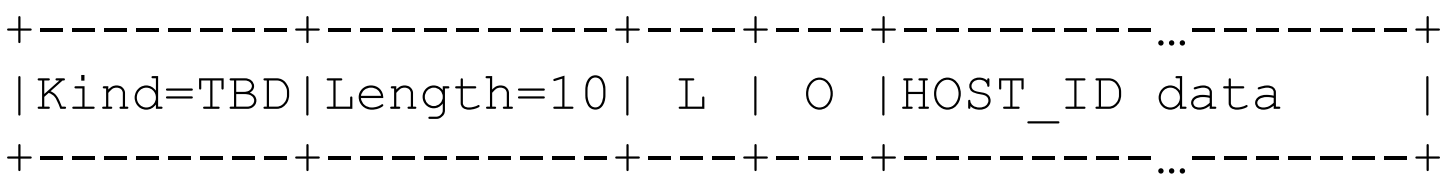
Appendix

HOST_ID_WING

HOST_ID_WING is sent in the SYN packet

```
+---------+--------+--------------+
|Kind=TBD |Length=4| HOST_ID data |
+---------+--------+--------------+
HOST_ID data: 16 bits
HOST_ID data can be:
   - lower 16 bits of the IP address
   - VLAN ID
   - VRF ID…
```

HOST_ID_BOUCADAIR



L: Lifetime (value=validity time; RFC6250)

0: Permanent

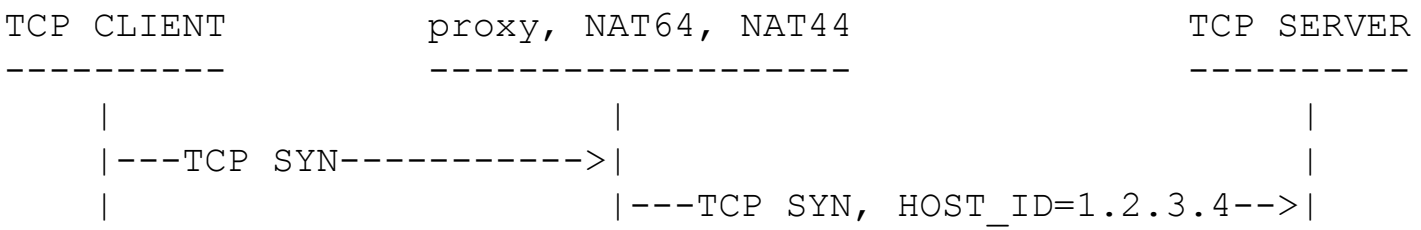
Origin:

- 0: Internal Port
- 1: Internal IPv4 address
- 2: Internal Port:Internal IPv4 address
- 3: IPv6 Prefix
- Else: No particular semantic;

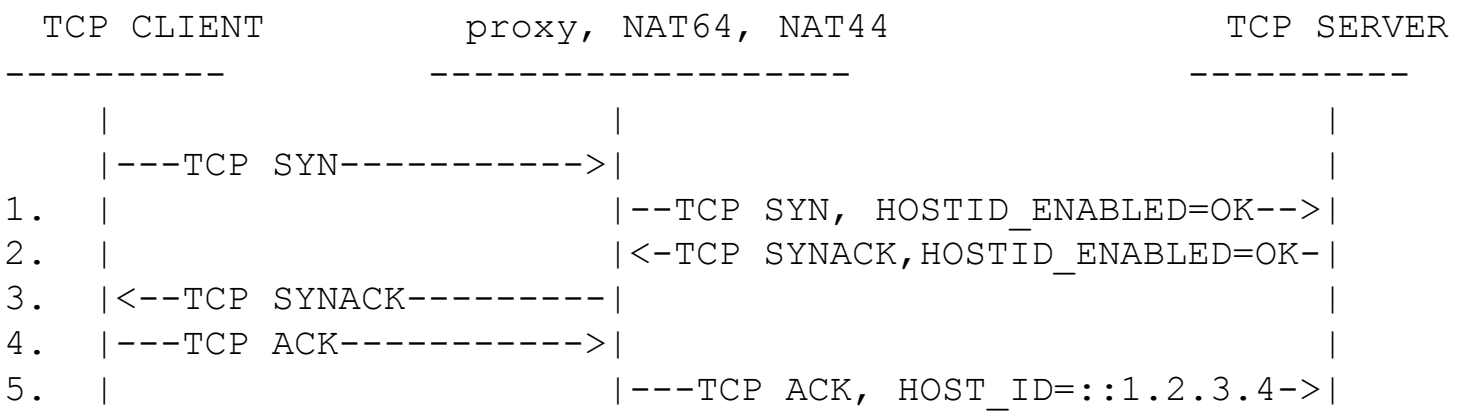
HOST_ID: depends on the content of the Origin field; padding is required

HOST_ID_BOUCADAIR

1. SYN Mode: the option is sent in the SYN packet



2. ACK Mode:
   1) Send HOST_ID_ENABLED in SYN
   2) If the remote TCP server supports that option, it must return it in SYNACK
   3) Then the TCP Client sends HOST_ID_BOUCADAIR in ACK



HTTP Results

No Impact for HOST_ID options on TCP session establishment delays

Alexa top 100,000 HTTP sites

	No-Option	O-WING	Difference NoOpt-WING	O-BOUCADAIR	Difference NoOpt-BOUCADAIR
1-1000	995	995	0	995	0
1001-2000	992	991	1	991	1
2001-3000	986	986	0	986	0
3001-4000	991	990	1	990	1
4001-5000	993	993	0	993	0
5001-6000	996	996	0	996	0
6001-7000	995	994	1	994	1
7001-8000	984	983	1	983	1
8001-9000	993	993	0	992	1
9001-10000	991	991	0	991	0
10001-20000	9785	9776	9	9776	9
20001-30000	9764	9747	17	9746	18
30001-40000	9778	9768	10	9766	12
40001-50000	9757	9746	11	9746	11
50001-60000	9771	9761	10	9761	10
60001-70000	9761	9752	9	9751	10
70001-80000	9744	9737	7	9736	8
80001-90000	9739	9730	9	9730	9
90001-100000	9736	9719	17	9719	17
1-100000	**97751**	**97648**	**103**	**97642**	**109**

No Impact for the Top1000 websites

Failure Ratio 0.105% for HOST_ID_*WING*

Failure Ratio 0.112% for HOST_ID_*BOUCADAIR*

6 HTTP servers did not respond HOST_ID_BOUCADAIR

FTP Results

- list from ftp-sites.org (5591 servers)
- 2045 FTP servers were reachable
- On average, no impact for HOST_ID options on TCP connection delays

	No-Option	O-WING	Failures	Failure Ratio
1-100	100	100	0	0,000%
101-200	100	99	1	1,000%
201-300	100	99	1	1,000%
301-400	100	100	0	0,000%
401-500	100	100	0	0,000%
501-600	100	100	0	0,000%
601-700	100	100	0	0,000%
701-800	100	100	0	0,000%
801-900	100	99	1	1,000%
901-1000	100	99	1	1,000%
1001-2000	1000	995	5	0,500%
2000-2045	45	45	0	0,000%
Total	**2045**	**2036**	**9**	**0,44%**

No Impact **for HOST_ID options on TCP session establishment delays**

Same Results for all HOST_ID options

Connection problems with *9 FTP servers for all HOST_ID options (0,44%)*

CGN (ISC-AFTR) Testing Results

N=10

		No-Option	O-WING	O-BOUCADAIR 3	O-ENABLED
TCP connection established		1378	1267	1363	1369
TCP SYN SENT		1378	1267	1363	1369
Success Ratio		100	100	100	100
TCP Retries		193	193	197	177
TCP timeouts		140	136	152	111
HTTP connection latencies	**t=20s**	**0,11**	**0,21**	**0,2**	**0,1**
	t=40s	**0,4**	**0,5**	**0,5**	**0,45**
	t=60s	**0,6**	**0,6**	**0,5**	**0,6**
HTTP throughput received (server)		46,47	45,31	45,88	46,12
TCP Connections Established/s(server)		20,29	19,88	20,06	20,18

Succes ratio is not impacted by HOST_ID options

No impact for HOST_ID options on **Connection Latencies**

N=5,000

		No-Option	O-WING	O-B1	O-B4	O-ENABLED
TCP connection established		1576	2000	1698	1796	1998
TCP SYN SENT		1794	2304	1980	2009	2262
Success Ratio		87	86	85	89	88
TCP Retries		3018	3101	2864	3013	3149
TCP timeouts		1167	1298	1064	1213	1417
HTTP Connection Latencies	**t=20s**	**2,2**	**3**	**1,4**	**2,2**	**2,5**
	t=40s	**3,7**	**3**	**3,1**	**3,3**	**3**
	t=60s	**7,8**	**5**	**6,3**	**7**	**5,6**
	t=70s	**9,6**	**6**	**7,4**	**8,7**	**7**
HTTP throughput received (kbps)(server)		45	54,52	48,65	51,45	57,2
TCP Connections Established/s (server)		19,8	24,05	21,45	22,45	25,05

Succes ratio is not impacted by HOST_ID options

No impact for HOST_ID options on **Connection Latencies**