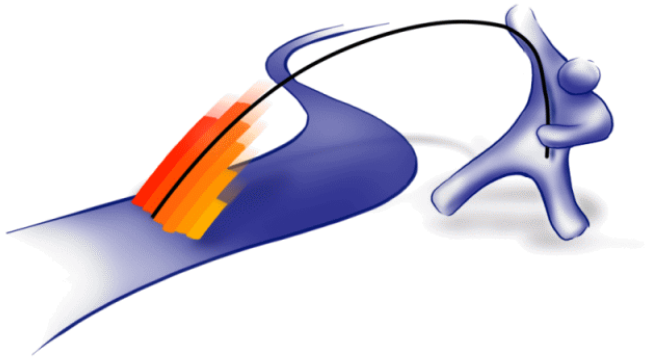
Carrier Grade NAT
Deployment, Traffic Impact, and Logging Considerations



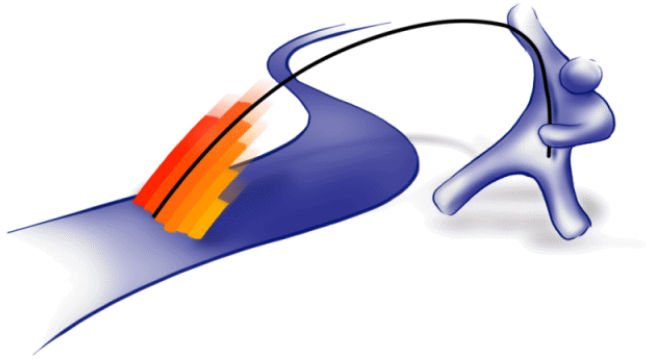
Jochen Kögel
IsarNet
jk@isarnet.de

13.03.2012

isarNet Software Solutions GmbH

Agenda

- Motivation
- NAT overview
- Implementation and Deployment
- NAT Logging
- Traffic impact
- Summary



Motivation

Problem

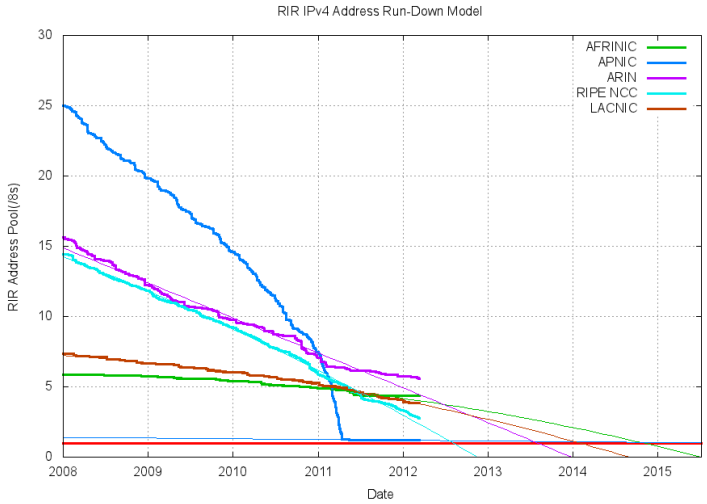
Increasing number of IP-enabled (mobile) devices

→ **IPv4 Address space exhausted**

Solutions

1. Efficient use of remaining IPv4 addresses
2. Migration to IPv6 (+translations to reach IPv4)

→ **both** require address translation!



http://www.potaroo.net/tools/ipv4/

Implementation

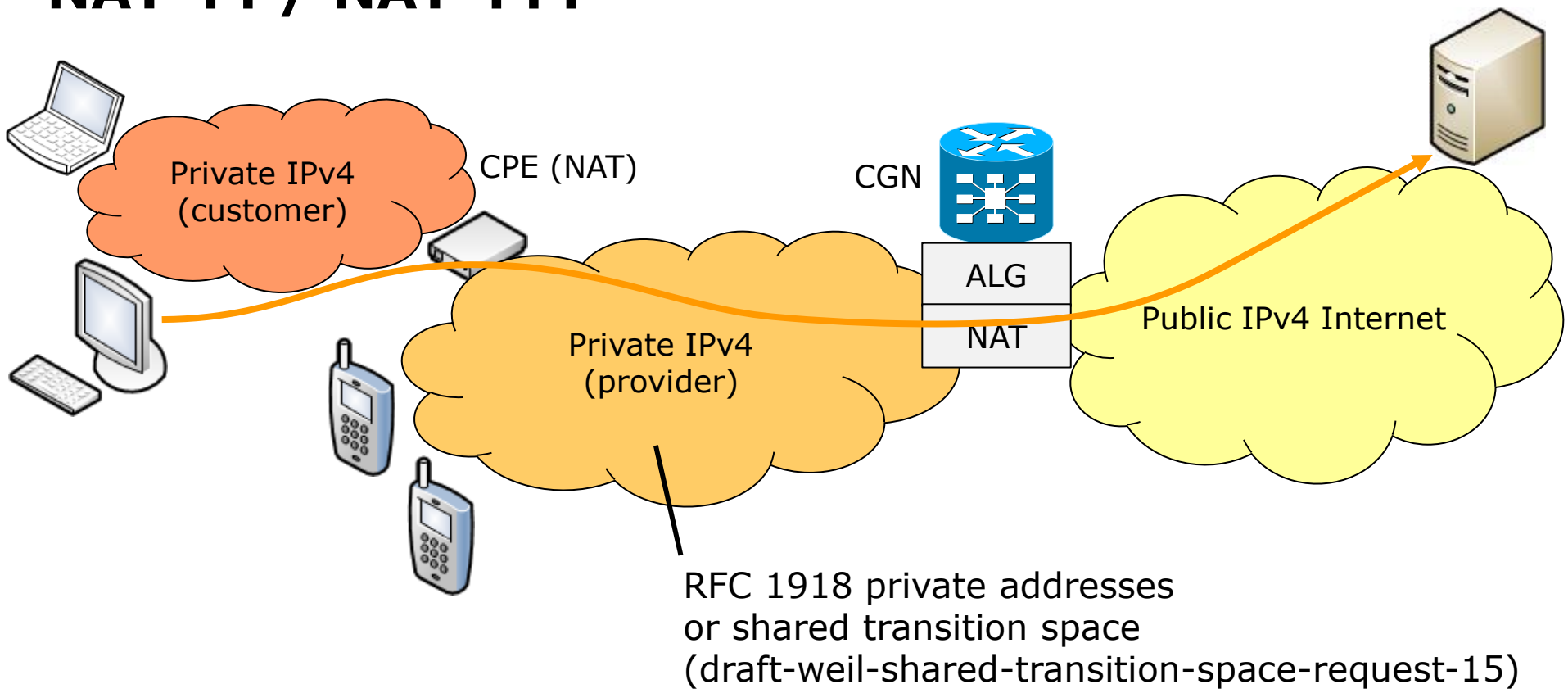
- Address translation **in the provider network**
- Carrier Grade Network Address Translation (**CGN**), also Large Scale NAT (**LSN**)

…CGNs will exist for several years (forever?)

…should use IPv4 addresses efficiently (many subscribers per public IPv4 Address)

NAT overview

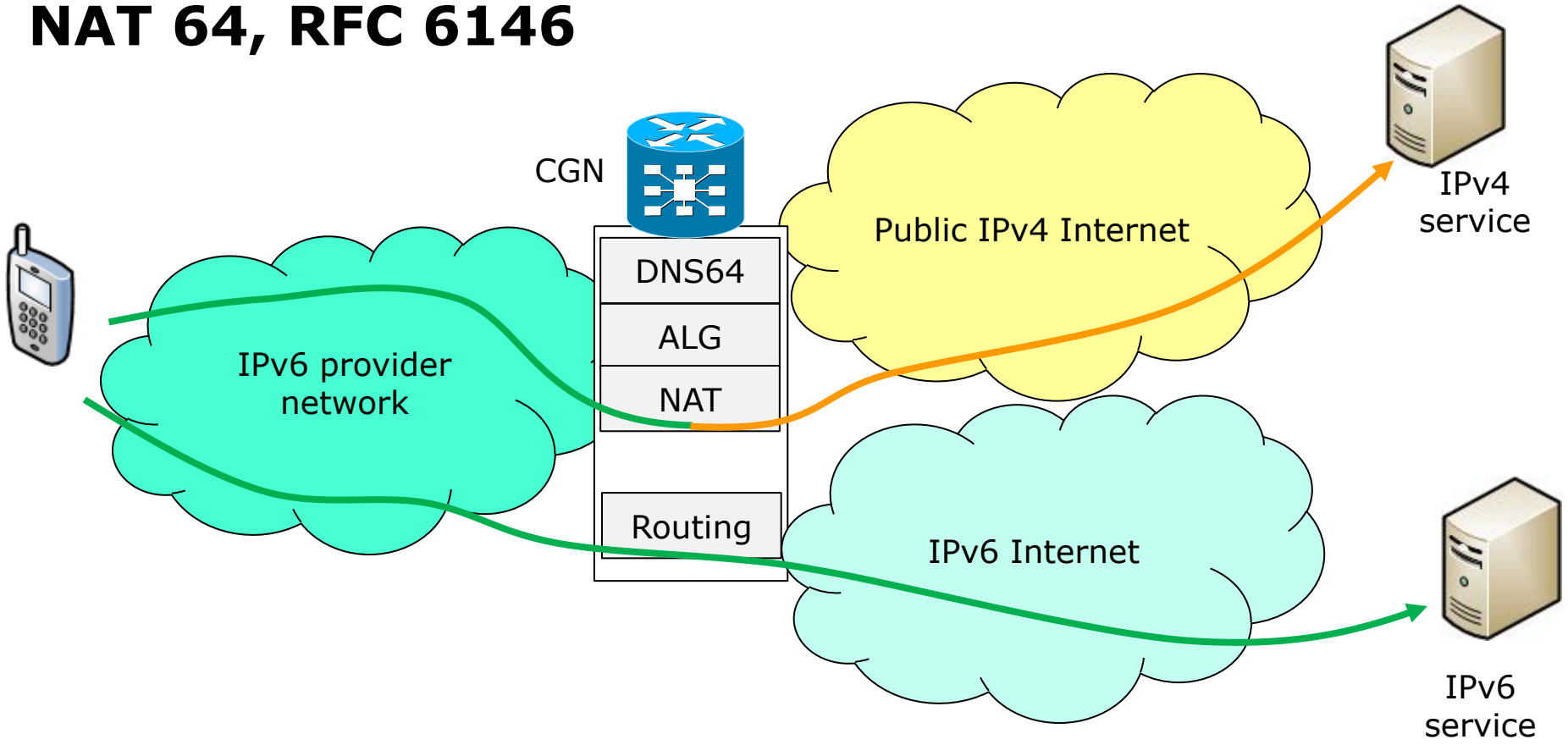
NAT 44 / NAT 444



- More efficient usage of IPv4 resources
- Application Layer Gateway (ALG) for IP-address-bound applications
- Short-term solution – no IPv6 deployed

NAT overview

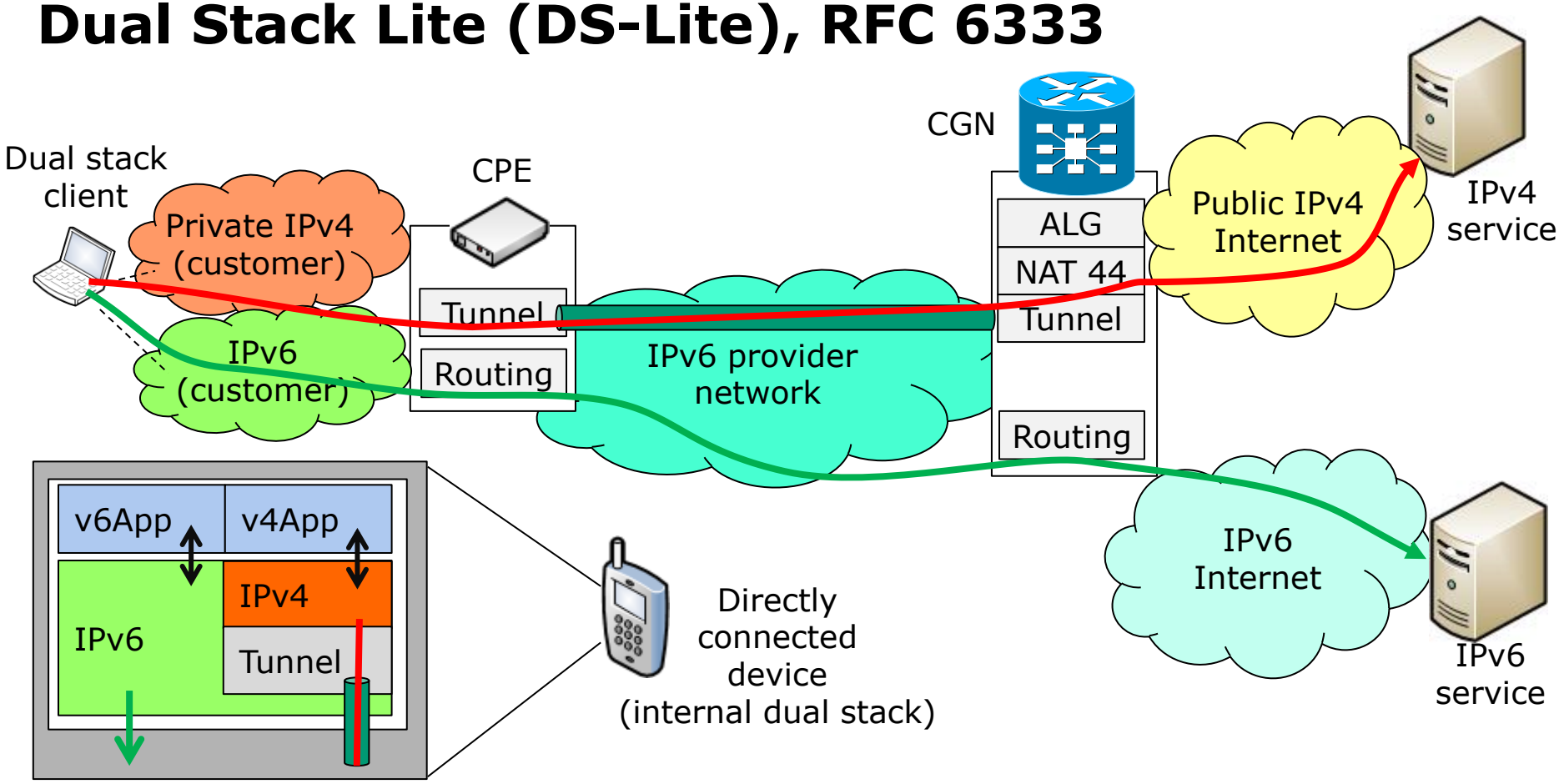
NAT 64, RFC 6146



- Provider network fully migrated to IPv6
- IPv6 clients only
- IPv4 content reachable via DNS64-translation and NAT
- Long-term solution – will be there for several years (?)

NAT overview

Dual Stack Lite (DS-Lite), RFC 6333

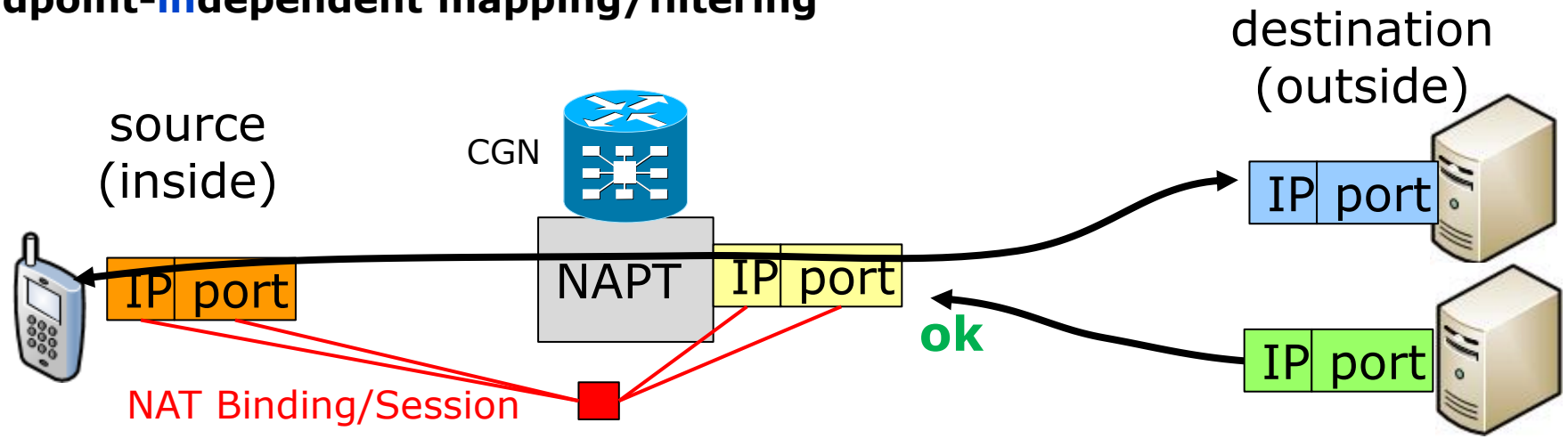


- Dual Stack from client point of view
  - IPv4 and/or IPv6 clients/applications
- Provider network IPv6 only

Implementation & Deployment

Network Address and Port Translation (NAPT) Implementations, RFC 4787

Endpoint-independent mapping/filtering

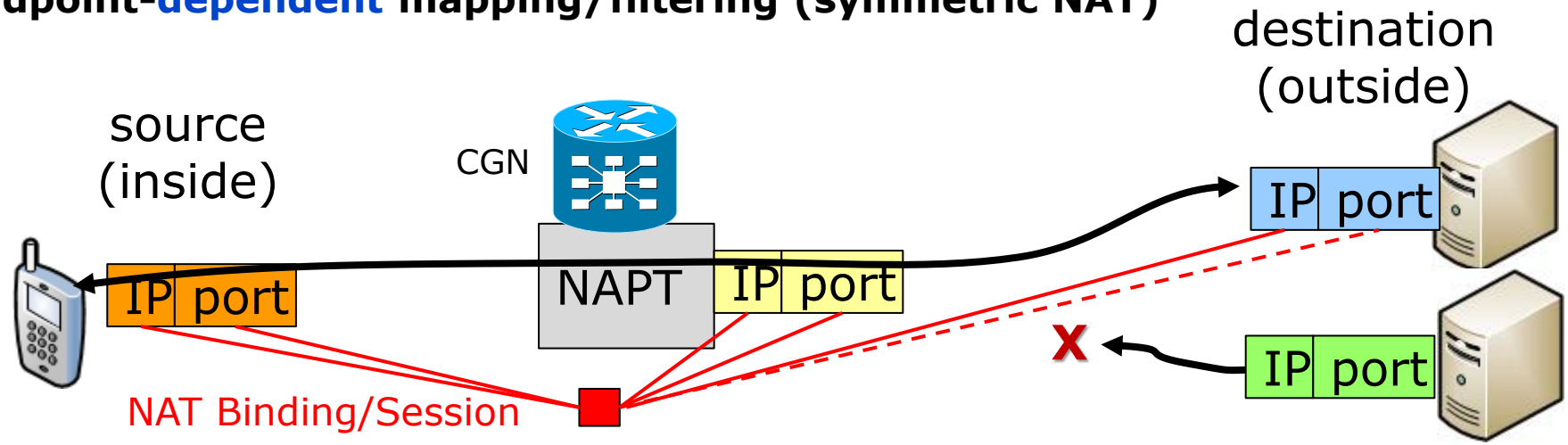


- Any endpoint can access client through NAT binding
- Prerequisite for NAT-traversal mechanisms (STUN, P2P applic., …)
- RFC 4787: MUST use endpoint-independent mapping

Implementation & Deployment

Network Address and Port Translation (NAPT) Implementations, RFC 4787

Endpoint-dependent mapping/filtering (symmetric NAT)



- Only initially addressed endpoint can access client through NAT binding
- Breaks NAT-traversal mechanisms (STUN, P2P, …) … use it deliberately?
- Current discussions: higher efficiency of endpoint-dependent mapping

Implementation & Deployment

Operational Requirements

- **NAT Efficiency**
  - Efficient resource (IPv4 address) usage
  - Free used resources as soon as possible
  - → **Timeouts** required
- **Security**
  - Several Address/port scans could exhaust public IP address pool easily
  - → Limit maximum resource usage: **portlimit**
- **Availability**
  - → **Failover** capabilities
- **Logging**
  - How to track down malicious activities?
  - Public IP Address is no longer a „temporary user identifier"
  - → Private IP ↔ public IP **mapping required** in logs

Implementation & Deployment

Exemplary CGN device: Cisco CGSE for CRS-1

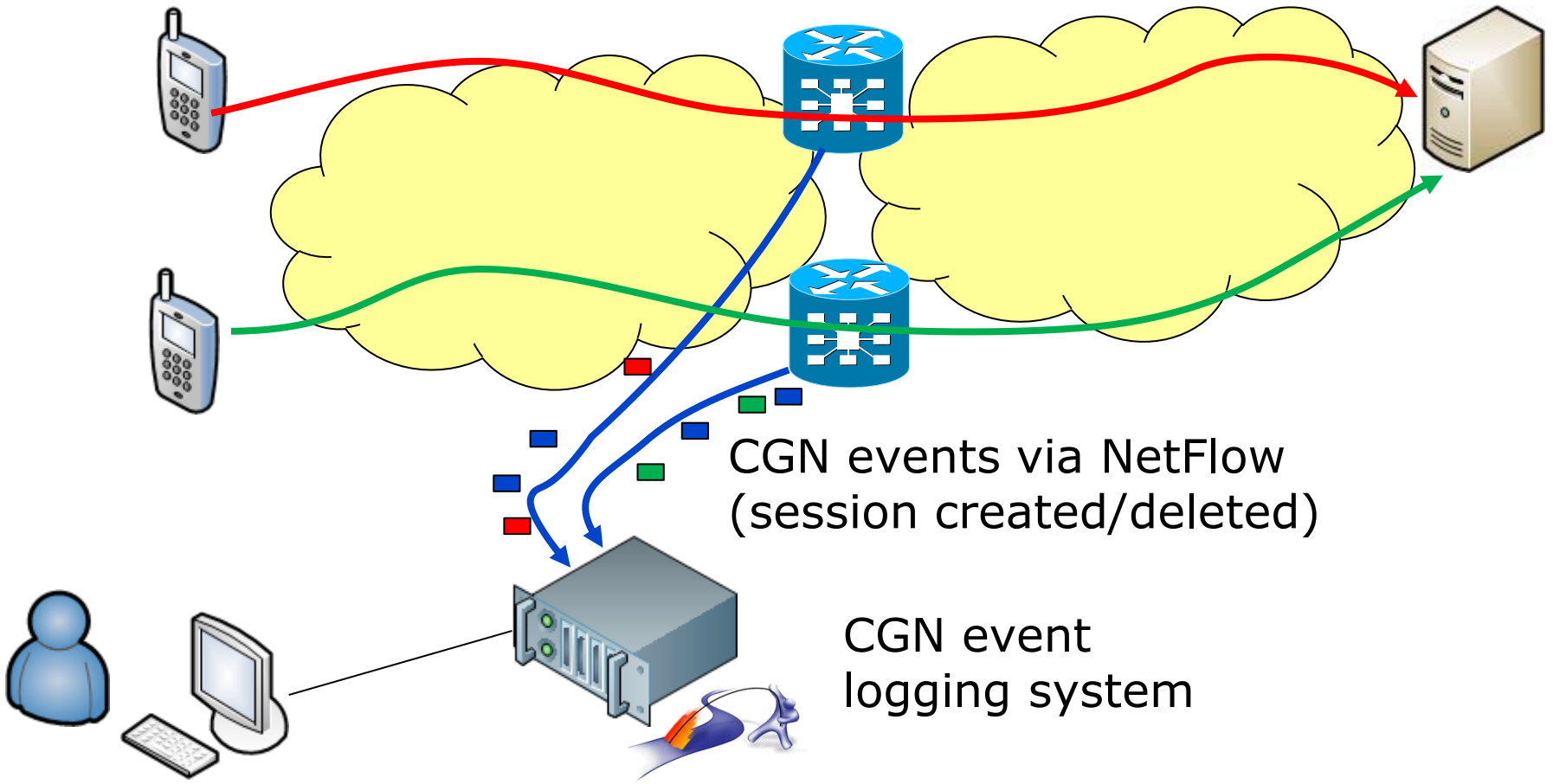
Performance

- Concurrent sessions: 20 Million
- Session creation rate: 1 Msessions/s
- Throughput: 20 Gbps full-duplex

Default configuration

- Portlimit 100
- Timeouts (inital/active)
  - TCP timeout 120s/1800s
  - UDP timeout 30s/120s
  - ICMP timeout: 60 s

NAT Logging


CGN events via NetFlow
(session created/deleted)
CGN event
logging system
Log analysis
Special version of IsarFlow
monitoring system:
IsarFlow CGN

NAT Logging

Logging requirements and solutions

Worst case scenario for dimensioning: failover

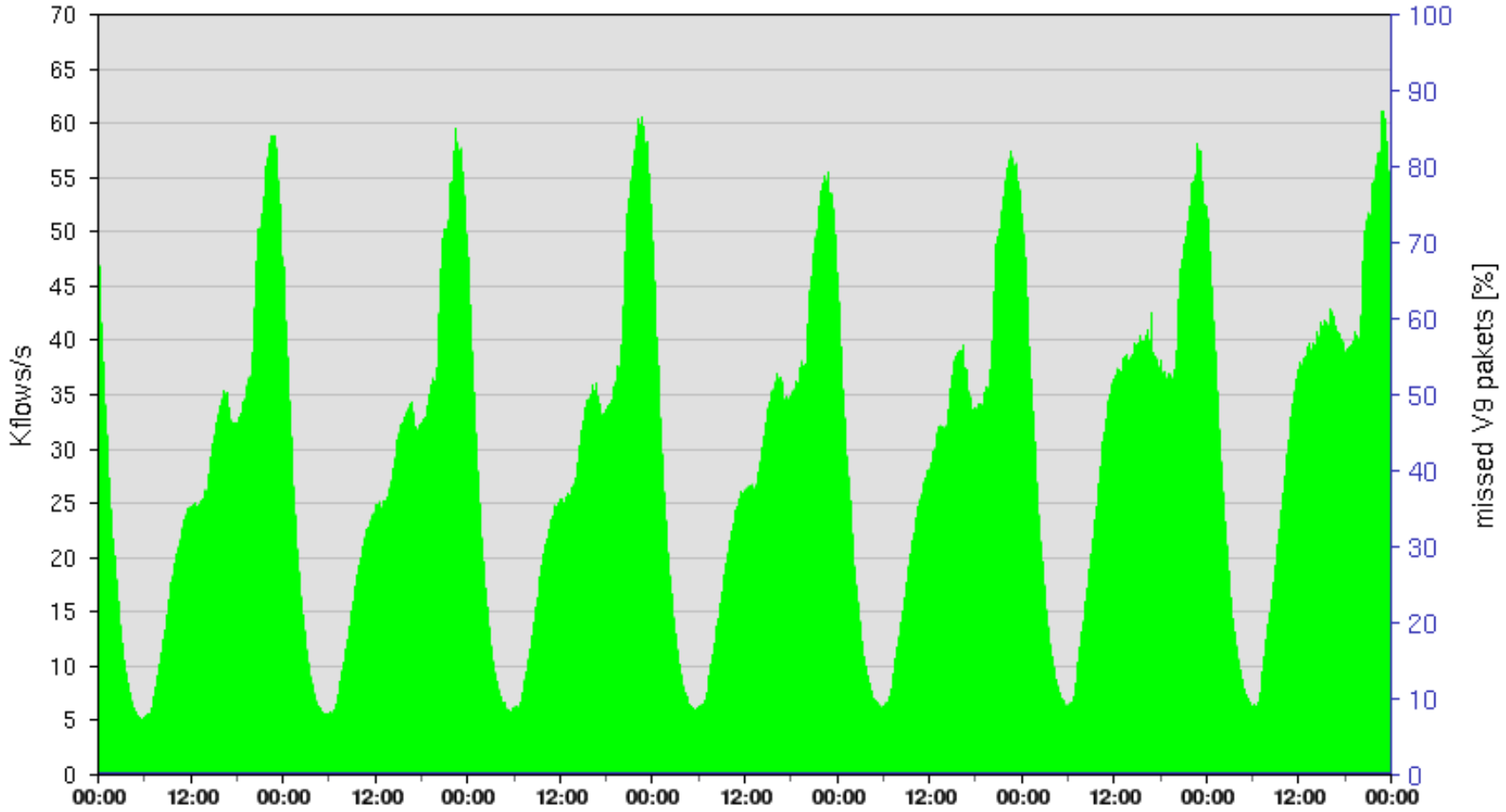
- All sessions of one location move to other locations
- Event-burst from one CGSE
  - 20 Million add events in 20 seconds
  - 1 Mevents/s → **180 Mbit/s**

Performance of IsarFlow CGN

- NAT 44 performance per server (COTS Linux box)
  - 1.5 Mevents/s → **270 Mbit/s** max. sustained rate
  - Loading of data into compressed database at that rate
  - Peak rate beyond 3 Mevents/s (540 Mbit/s) without loss
- Performance **scales with number of servers** (distributed DB)
- Storage requirements: **8 MB for 1 Million sessions**

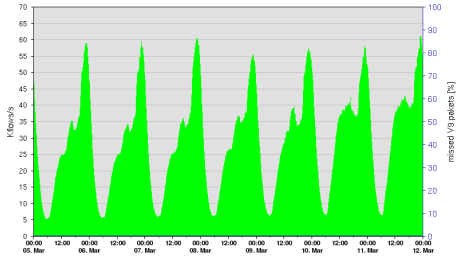
NAT Logging

Typical event rate patterns



NAT Logging

Ideas for reducing logging effort



Bulk port allocation

- Allocate several ports at once for each client

→ One log event for large port range

Problem: deterministic source ports are a security problem (RFC 6056)

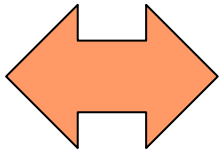
Possible solution: algorithmic port scattering

General problem of bulk port allocation

- Bulk allocation is „**port over-provisioning**" → lower NAT efficiency

Trade-off

NAT efficiency
(cost of public IPv4 addresses)

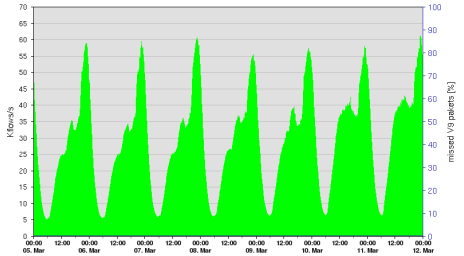


Logging efficiency
(cost of hard disk space)

Traffic impact

First studies of traffic characteristics

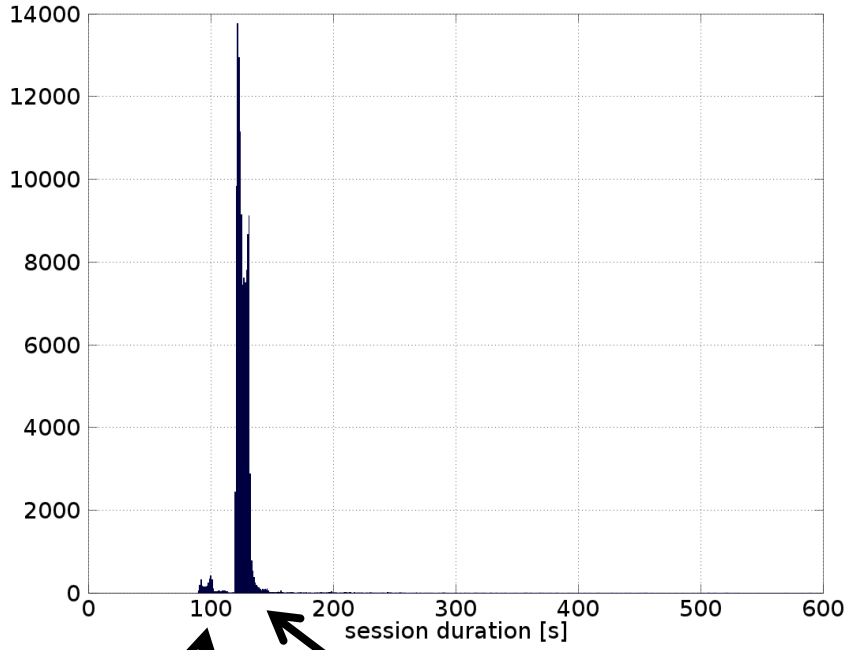
Based on inspection of 400k sessions



- No surprise: lots of **short sessions**
  - → Cause high NAT event rates
  - → Affect NAT efficiency: timeout until port can be reused
- Avg. Session duration (incl. timeouts) – 135 s
- > 30 % UDP sessions (**!**)
  - DNS Resolver in public IP space?
  - SIP/VoIP also across NAT?

Traffic impact

UDP



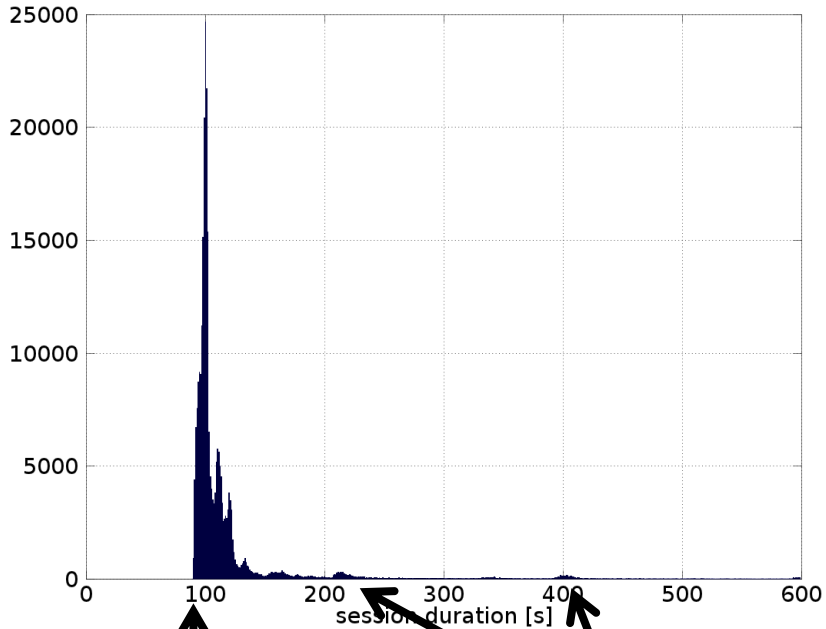
Failed UDP Requests (?)

1 packet
→ initial timeout

Successful UDP requests (?)

>1 packets
→ active timeout

TCP



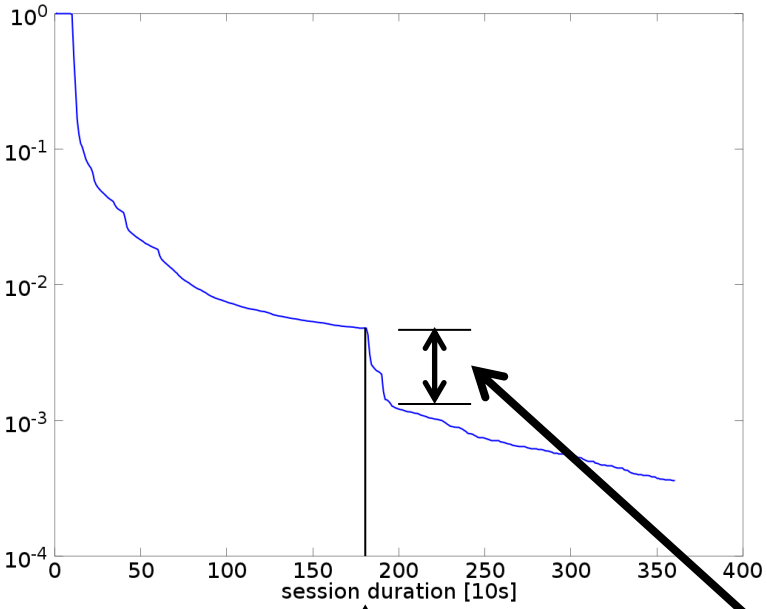
NAT timeout after connection termination

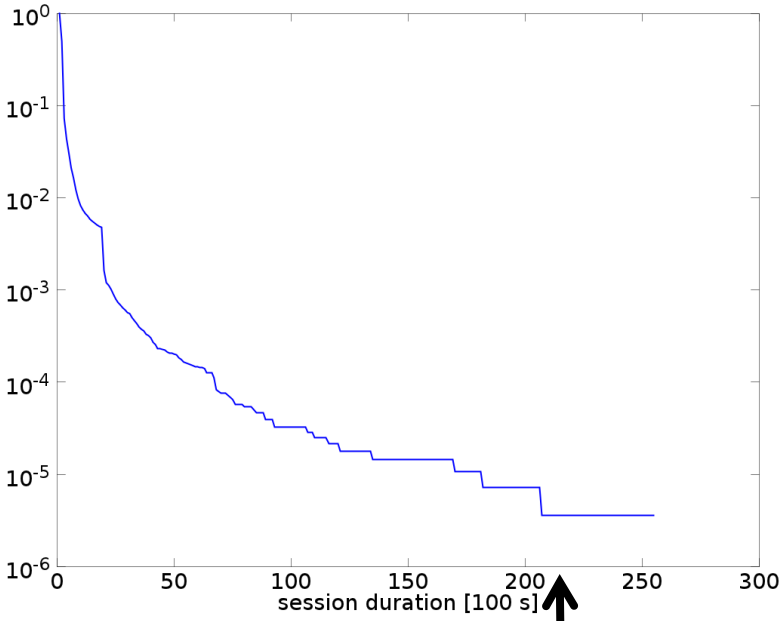
or after failed connection setup

Application timeouts (browser connections, …)

Traffic impact

TCP - CCDF





NAT TCP active session timeout

~ TCP connections that ran into active timeout (broken? Or have still been „alive“?)

6h session duration

Summary & Outlook

Summary

- Carrier grade NATs are currently getting deployed
- Still lots of standardization effort and new ideas (IETF Softwire, Behave, v6ops…)
- **NAT session logging** is a major concern
  - **IsarFlow CGN** proves feasibility of large-scale full session logging
  - Bulk allocation: Trade-off between NAT efficiency and logging
- Traffic impact
  - Portlimit: limited number of sessions per user
  - Timeouts: keepalives necessary, as with current CPE NATs
  - Short sessions: Cause high event rates and block resources due to timeouts

Outlook

- How will mobile session behavior evolve?
  - Impact on NAT efficiency
  - Impact on NAT configuration (timeouts, …)
  - Impact on Logging requirements


isarNet
Software Solutions GmbH

References

IETF documents

- RFC 4787: Network Address Translation (NAT) Behavioral Requirements for Unicast UDP
- draft-ietf-behave-lsn-requirements
- RFC 6146: Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers
- RFC 6535: Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)
- RFC 6144: Framework for IPv4/IPv6 Translation
- RFC 6052: IPv6 Addressing of IPv4/IPv6 Translators
- RFC 5382: NAT Behavioral Requirements for TCP
- RFC 6056: Recommendations for Transport-Protocol Port Randomization
- RFC 6333: Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion
- RFC 6269: Issues with IP Address Sharing