



forthnet



RIPE 67
ATHENS / GREECE

Moving on...

IPv6, DS-Lite & PCP



Anastasios Chatzithomaoglou
(achatz@forthnet.gr)

IP Engineering – Forthnet

17-10-2013

IPv6

IPv6 in www.forthnet.gr

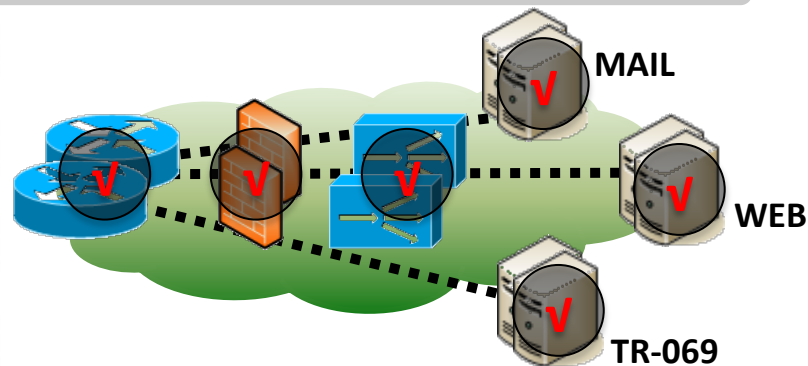
...in just a few hours...

Reserve IPv6 prefixes

Configure IPv6 prefixes

Update IPv6 filters

Update DNS



Forward Reload Stop

http://www.forthnet.gr/Services3PlayFormJul13.aspx?a_id=6208

6⁴

http://www.forthnet.gr	2a02:2148:99::80
http://platform.twitter.com	68.232.35.139
http://lp.longtailvideo.com	93.184.221.48
https://cdn.api.twitter.com	2.23.113.224

FORTHNET.GR NOVASPORTS.GR NOVACINEMA.GR DRIVEME.GR OPENSEAS.GR ΤΗΛΕΟΠΤΙΚΗ

HOME Ο Όμιλος ΥΠΗΡΕΣΙΕΣ ΓΙΑ ΤΟ ΣΠΙΤΙ ΥΠΗΡΕΣΙΕΣ ΓΙΑ ΤΗΝ ΕΠΙΧΕΙΡΗΣΗ ΕΞΥΠΗΡΕΤΗΣΗ

forthnet nova

Επικοινωνία

Υπηρεσίες για το Σπίτι

ΕΠΙΛΕΞΤΕ: 3play | 2play | Internet | Τηλεφωνία | Τηλεόραση

Statistics

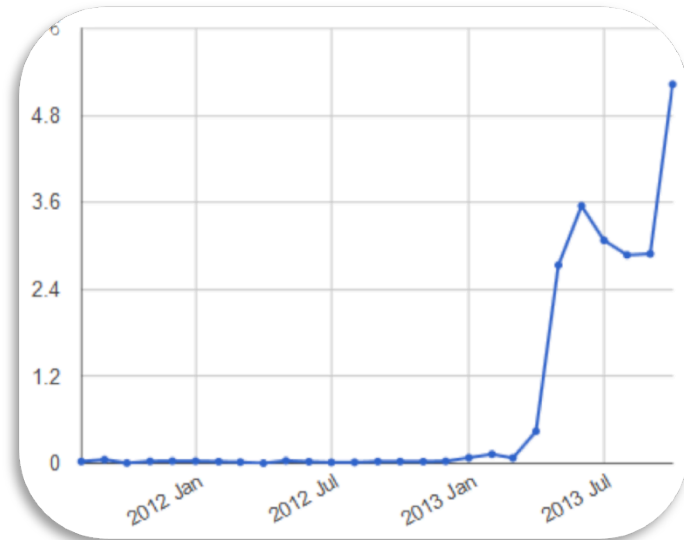
IPv6

IPv6 measurements

RIPE: 4,8%

APNIC: 5%

Access (last 6 months)	Access (last month)	Content	LIR
3.7 %	4.8 %		FORTHnet SA



W6L: 2,35%

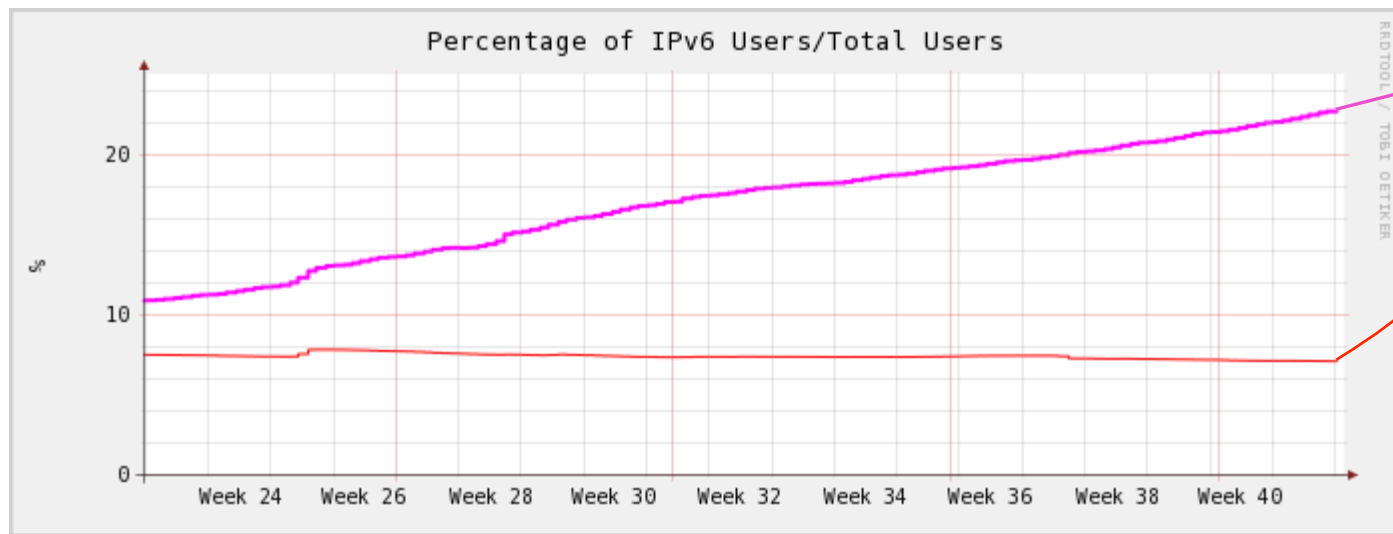
Participating Network	ASN(s)	IPv6 deployment
Forthnet	1241	2.35%

IPv6

IPv6 subscribers

More than **10%** of subscribers could access the internet over IPv6 in **June**

Almost **23%** of subscribers should access the internet over IPv6 in **October**

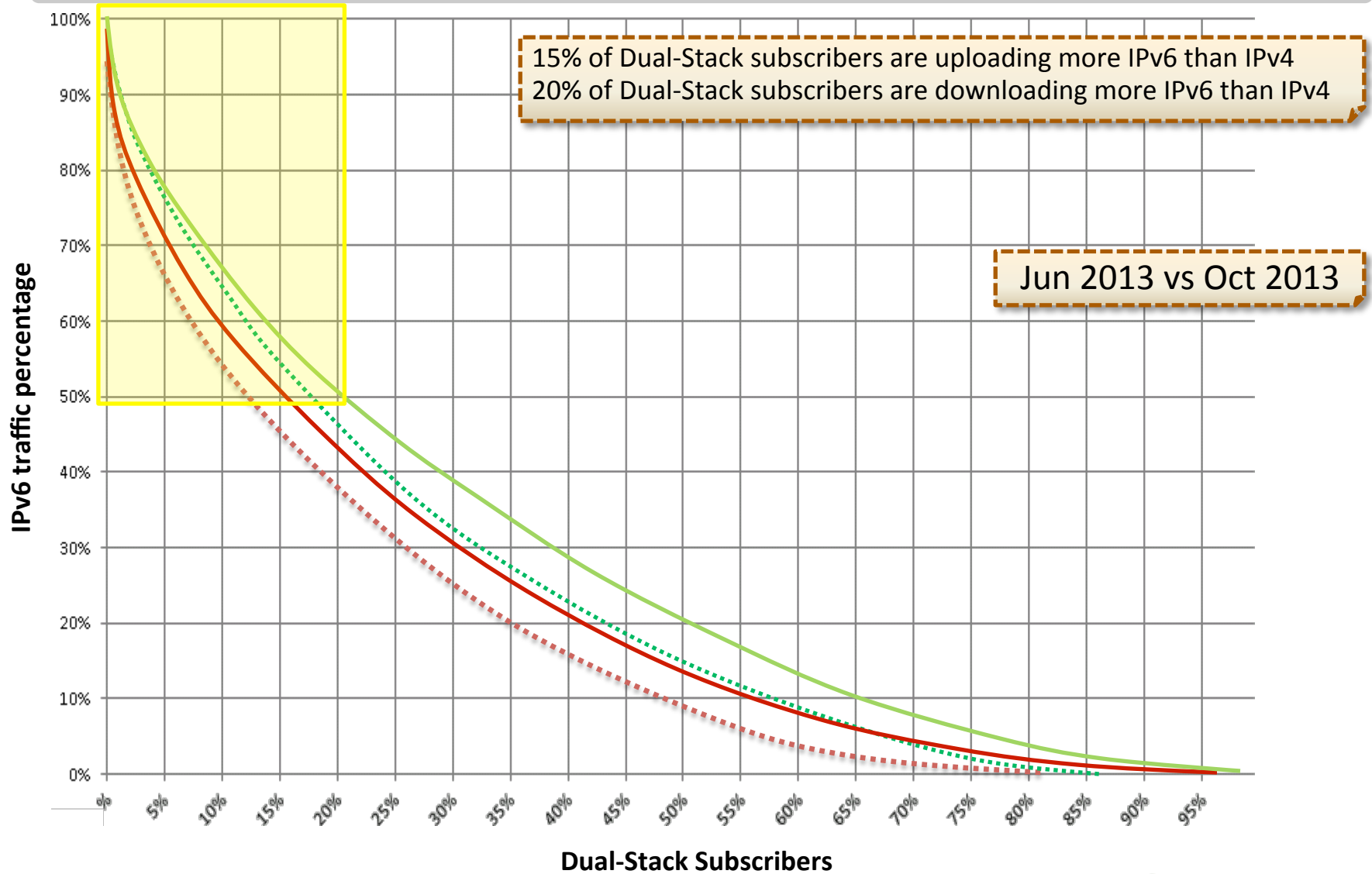


But only **8%** can actually do it
Why?

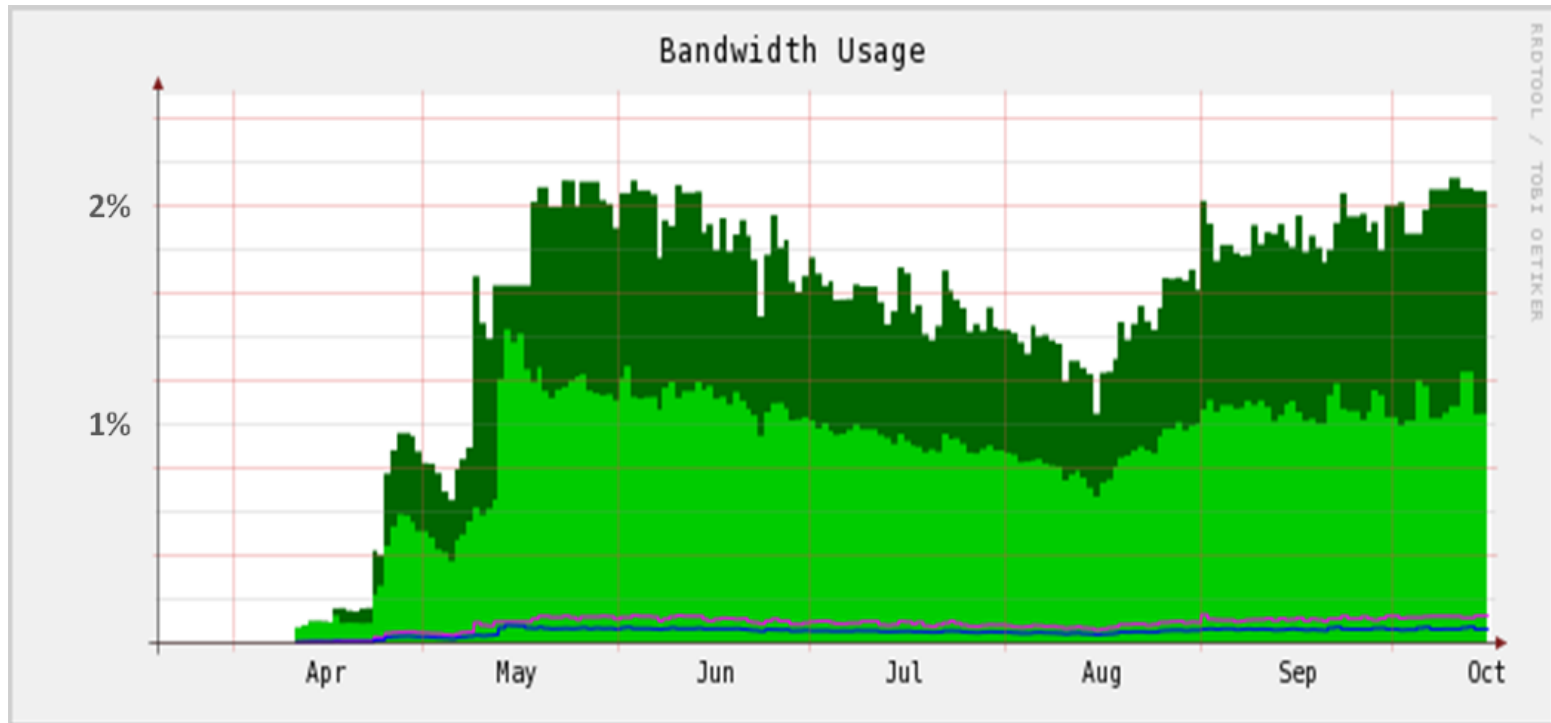
A wrong default value in the latest
firmware of one of our CPEs

IPv6

IPv6 traffic vs IPv4 traffic



Almost **2%** of our traffic is over IPv6



We expect that to reach **6%** in 2014

IPv6 in CPEs

- ❑ 5 CPEs with Dual-Stack (6 firmwares avg)
- ❑ 5 CPEs with DS-Lite (3 firmwares avg)
- ❑ 1 CPE with PCP (5 firmwares avg)

Most common problems

- ❑ DHCPv6-PD not working
- ❑ Passing DNS info to clients
- ❑ Non-optimal MTU

IPv4 Address Depletion

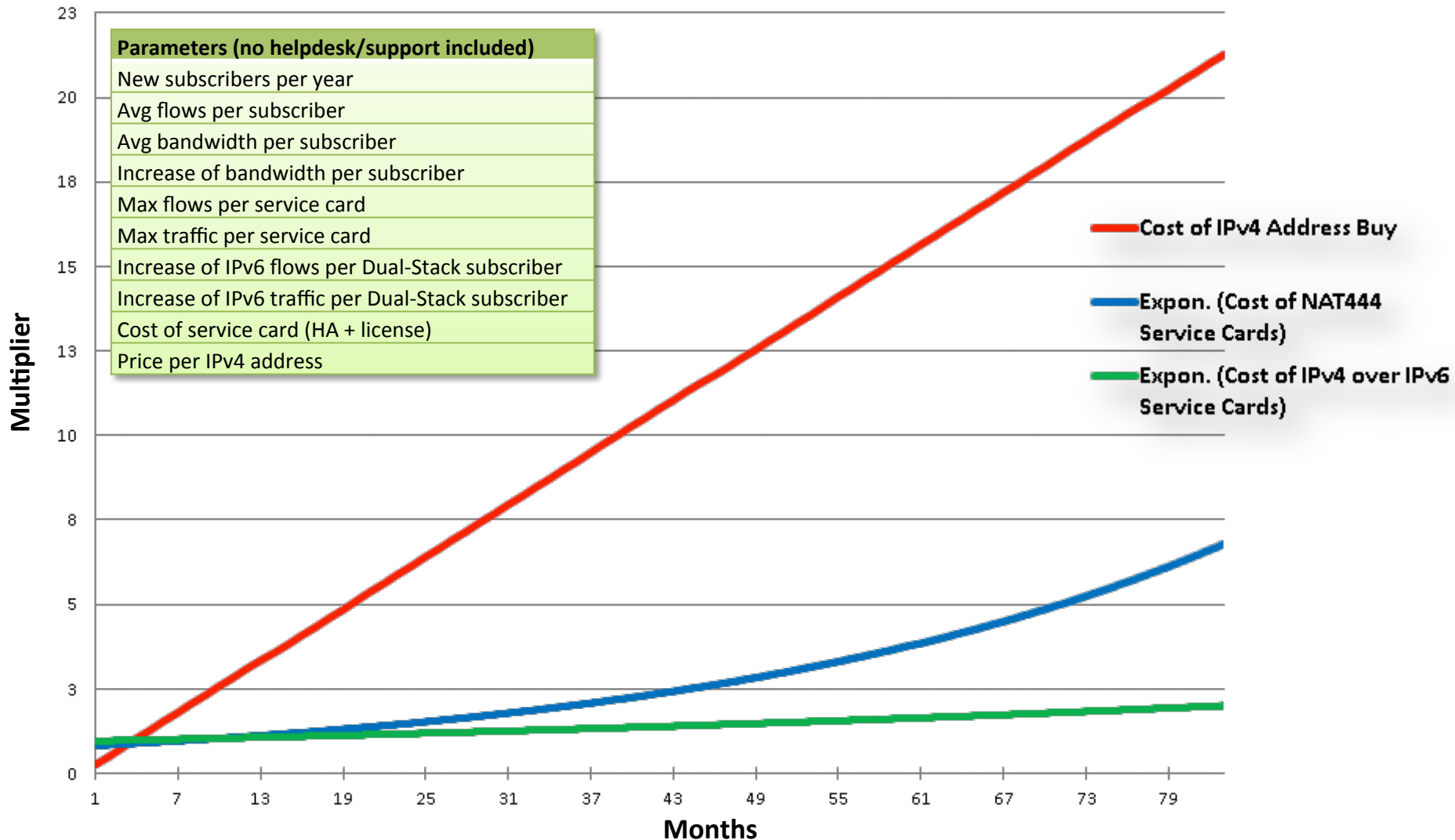


DS-Lite

Migrating from Decentralized IPv4 Assignment to Centralized

- ❑ Currently each BRAS has its own IPv4 local pool (distributed model)
- ❑ There exist 5 BRAS clusters (to become 3 soon), each one with an overflow BRAS
- ❑ Due to BRAS clustering, only the overflow BRAS has underutilization of its local IPv4 address pool
- ❑ A migration is underway in order to move distributed pools from overflow BRAS of all clusters to a centralized DHCPv4 server
- ❑ Routing de-aggregation will be minimal
- ❑ An **extension of a few months** is expected for the IPv4 depletion date

NAT444 vs DS-Lite Cost



DS-Lite (Dual-Stack Lite) is a well-known **IPv6 transition technology**, while at the same time offers a **solution for the depletion of IPv4 addresses**.

The underlying architecture of the DS-Lite uses only IPv6 communication between the provider and the subscriber's CPE, while retaining the IPv4 (or dual-stack) subscriber devices as they are.

The main reasons for selecting DS-Lite for Forthnet's network are:

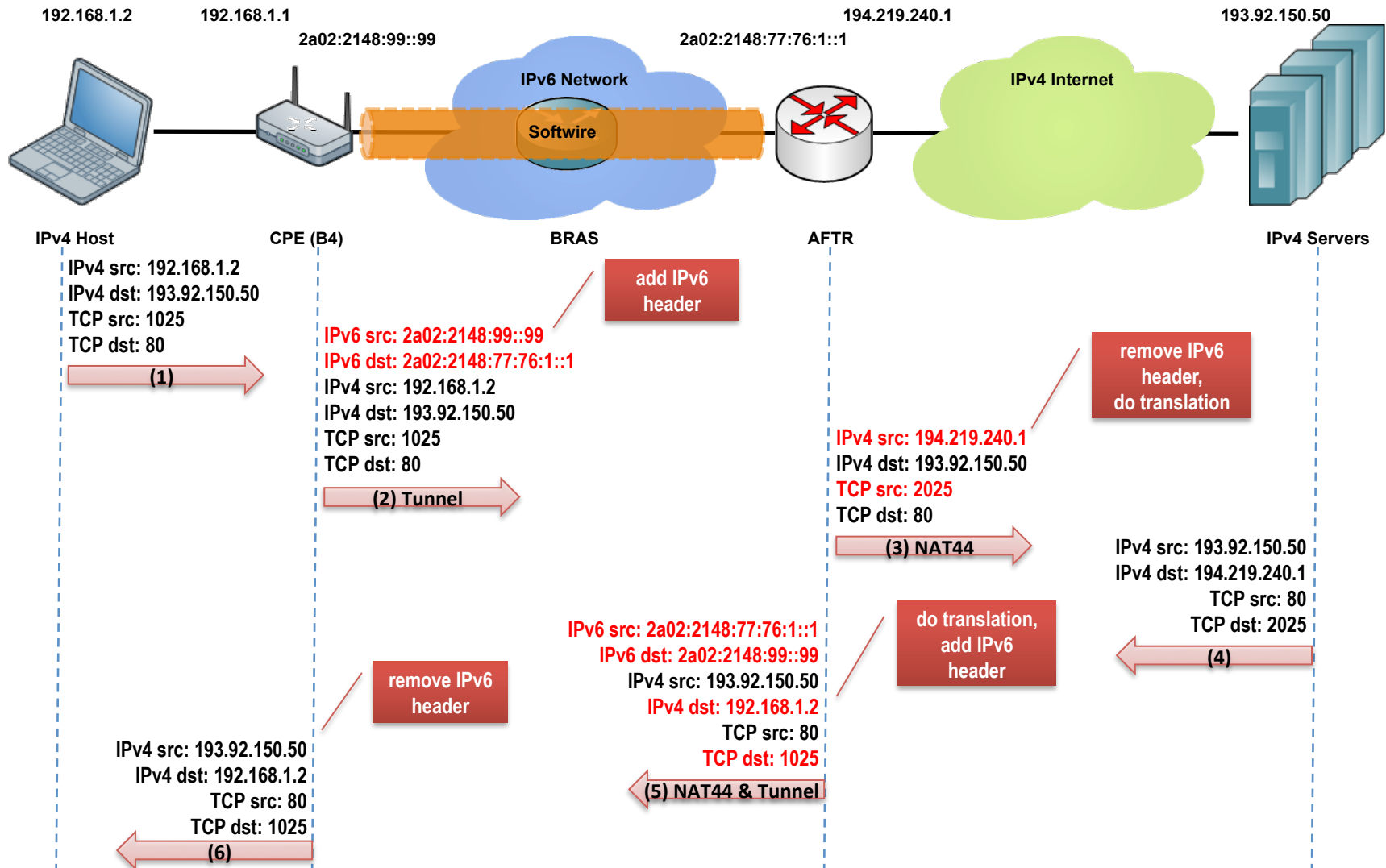
- the **mature** technology
- the **easy** implementation and provisioning
- **the use of IPv6 as its basis**

It is considered a technology solution with a relatively limited lifetime, since as IPv6 content/traffic increases, the need for communication via IPv4 will gradually decrease.

A public pilot is active in Forthnet's network since 1-Apr-2013.

Forthnet & DS-Lite

DS-Lite Operation



Current Logging Parameters

- ☐ Date/Time
- ☐ Reporting device identifier
- ☐ Subscriber identifier (Username/Circuit-ID)
- ☐ IPv4 address/network
- ☐ WAN IPv6 Prefix (used for correlation with the new logging parameters)
- ☐ LAN IPv6 Prefix

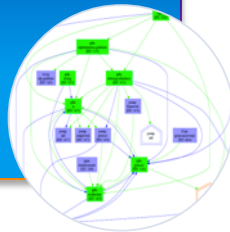
New Logging Parameters

- ☐ Reporting device type (Optional)
- ☐ Reporting device identifier (Optional/Mandatory per case)
- ☐ Subscriber identifier (Mandatory)
- ☐ External IPv4 address (Mandatory)
- ☐ External port or ICMP identifier (Mandatory)
- ☐ Protocol identifier (Optional/Mandatory per case)
- ☐ Internal port or ICMP identifier (Optional/Mandatory per case)
- ☐ Internal address of the source of the packet (Optional/Mandatory per case)

Dst IPv4 Address is not required/
recommended due to privacy issues
Dst owner should keep track of source ports

- efficient
- customized
- requires collector
- requires extra application

IPFIX



- readable
- searchable
- reduced performance
- larger volume

SYSLOG



SYSLOG could be used on a back-end device processing IPFIX records in real-time

Sample output from AFTR

Original Message	Stored Message		
2013 Sep 14 19:19:17.698 aftr-klm-01_re0 (FPC Slot 2, PIC Slot 0) 2013-09-14 16:19:17 {DSLITE-SERVICE-SET}[FWNAT]: ASP_SFW_CREATE_ACCEPT_FLOW: proto 6 (TCP) application: tcp, xe-0/1/0.0:192.168.1.5:65412 -> 111.111.111.111:80, creating forward or watch flow ; source address and port translate to 194.219.240.52:1886 ;Softwire 2a02:2148:100:301:f53a:a74a:1c28:a36f->2a02:2148:77:76:1:0:0:1	Field	Original Value	Final Value
	Date/Time	2013 Sep 14 19:19:17	19:19:17
	Action	Create/Delete	C/D
	Protocol	TCP	6
	Src Int IPv4 Address	192.168.1.5	-
	Src Int IPv4 Port	65412	-
	Src Ext IPv4 Address	194.219.240.52	194.219.240.52
2013 Sep 14 19:29:19.537 aftr-klm-01_re0 (FPC Slot 2, PIC Slot 0) 2013-09-14 16:29:19 {DSLITE-SERVICE-SET}[FWNAT]: ASP_SFW_DELETE_FLOW: proto 6 (TCP) application: any, (null)(null)192.168.1.5:65412 -> 111.111.111.111:80, deleting forward or watch flow ; source address and port translate to 194.219.240.52:1886 ;Softwire 2a02:2148:100:301:f53a:a74a:1c28:a36f->2a02:2148:77:76:1:0:0:1	Src Ext IPv4 Port	1886	1886
	Dst IPv4 Address	-	-
	Dst IPv4 Port	-	-
	Src IPv6 Address	2a02:2148:100:301:f53a:a74a:1c28:a36f	2a02:2148:100:301
	Total Bytes	114	56

Date is not required, because it's embedded in each log filename

Dst IPv4 Address is not required/recommended (Dst owner should keep track of source ports too)

Further decrease can be accomplished by omitting some delimiters, prefixes and the Delete/Release actions

Logging Volume Calculations

An average subscriber generates 27k connections per 24 hours (assuming not idle)

An bittorent user makes 27k connections per hour! 1/5 of these are unique...

Logging messages can have a length of 114 bytes or 56 bytes if optimized

- ❑ $27k \times 114 = 3,1 \text{ MB}$ logs per user per day

- ❑ $27k \times 56 = 1,5 \text{ MB}$ logs per user per day if optimized

A CGN can serve at least 32k subscribers

- ❑ $32k \times 3,1 \text{ MB} = 99 \text{ GB}$ per day, aka 3,0 TB per month

- ❑ $32k \times 1,5 \text{ MB} = 48 \text{ GB}$ per day, aka **1,4 TB per month** if optimized

Assumptions

- ❑ No compression is happening

 - ❑ Can provide 1/10 smaller size

 - ❑ Makes searching harder and slower

- ❑ No traffic is switched over to IPv6

 - ❑ IPv6 traffic is currently 2% of IPv4, showing exponential increase

 - ❑ New IPv4 subscribers are added every day

Sample output from AFTR (PBA+PCP)

Original Message	Stored Message		
2013 Sep 19 10:08:10.850 aftr-klm-01_re0 (FPC Slot 8, PIC Slot 0) 2013-09-19 07:08:10 [FWNAT]:ASP_NAT_PORT_BLOCK_ALLOC: 2a02:2149:ff01:6869:7c84:6829:ca04:f3cd -> 194.219.240.252:5120-6143	Field	Original Value	Final Value
	Date/Time	2013 Sep 19 10:08:10	10:08:10
	Action	Create/Delete Allocate/Release	C/D/A/R
2013 Sep 19 09:58:48.602 aftr-klm-01_re0 (FPC Slot 8, PIC Slot 0) 2013-09-19 06:58:48 {TEST-PCP-SERVICE-SET} [FWNAT]:ASP_PCP_NAT_MAP_CREATE: 2a02:2149:ff01:6869:7c84:6829:ca04:f3cd 192.168.1.2:48000 -> 194.219.240.252:48000	Protocol	?	?
	Src Int IPv4 Address	192.168.1.2	-
	Src Int IPv4 Port	48000	-
2013 Sep 19 10:50:59.933 aftr-klm-01_re0 (FPC Slot 8, PIC Slot 0) 2013-09-19 07:50:59 [FWNAT]:ASP_NAT_PORT_BLOCK_RELEASE: 2a02:2149:ff01:6869:7c84:6829:ca04:f3cd -> 194.219.240.252:3072-4095 0x523aa078	Src Ext IPv4 Address	194.219.240.52	194.219.240.52
	Src Ext IPv4 Port Range	48000 5120-6143	48000 5120-6143
	Dst IPv4 Address	-	-
	Dst IPv4 Port	-	-
	Src IPv6 Address	2a02:2149:ff01:6869: 7c84:6829:ca04:f3cd	2a02:2149:ff01:6869
2013 Sep 19 10:58:48.924 aftr-klm-01_re0 (FPC Slot 8, PIC Slot 0) 2013-09-19 07:58:48 {TEST-PCP-SERVICE-SET} [FWNAT]:ASP_PCP_NAT_MAP_DELETE: 2a02:2149:ff01:6869:7c84:6829:ca04:f3cd 192.168.1.2:48000 -> 194.219.240.252:48000	Total Bytes	124	62

Date is not required, because it's embedded in each log filename

Dst IPv4 Address is not required/recommended (Dst owner should keep track of source ports too)

Further decrease can be accomplished by omitting some delimiters, prefixes and the Delete/Release actions

PBA Logging Volume Calculations

An average subscriber makes 60 block allocations per 24 hours (assuming not idle)

An bittorent user makes 80 block allocations per 24 hours. Very little difference...

Logging messages can have a length of 124 bytes or 62 bytes if optimized

❑ $60 \times 124 = 7 \text{ KB}$ logs per user per day

❑ $60 \times 62 = 3,7 \text{ KB}$ logs per user per day if optimized

A CGN can serve at least 32k subscribers

❑ $32k \times 7 \text{ KB} = 238 \text{ MB}$ per day, aka 7 GB per month

❑ $32k \times 3,7 \text{ KB} = 119 \text{ MB}$ per day, aka **3,5 GB per month** if optimized

Outcomes & Comments

❑ No compression needed

❑ No optimization needed

❑ Little dependency on IPv6 traffic switchover

❑ PCP mappings make very little difference

❑ Depending on port block size and timeout, further optimization can be achieved

Next steps regarding IPv4 depletion

- ☐ Focus on technologies that lead to logging per port-set (natx4-log-reduction, deterministic-cgn, etc.)
- ☐ Evaluate Lw4over6 (move NAT from AFTR to B4)
- ☐ Evaluate other stateless solutions like MAP (still lacking CPE vendor interest)
- ☐ Implement automatic correlation of BRAS radius acct records with AFTR NAT mapping logs, unless RFC 6736 gets implemented first

Migrating customers to native
IPv6 is always **top priority**

PCP

PCP

- ❑ Port Control Protocol
- ❑ RFC 6887
- ❑ Server & Client/Proxy
- ❑ UPnP IGD-PCP Interworking Function

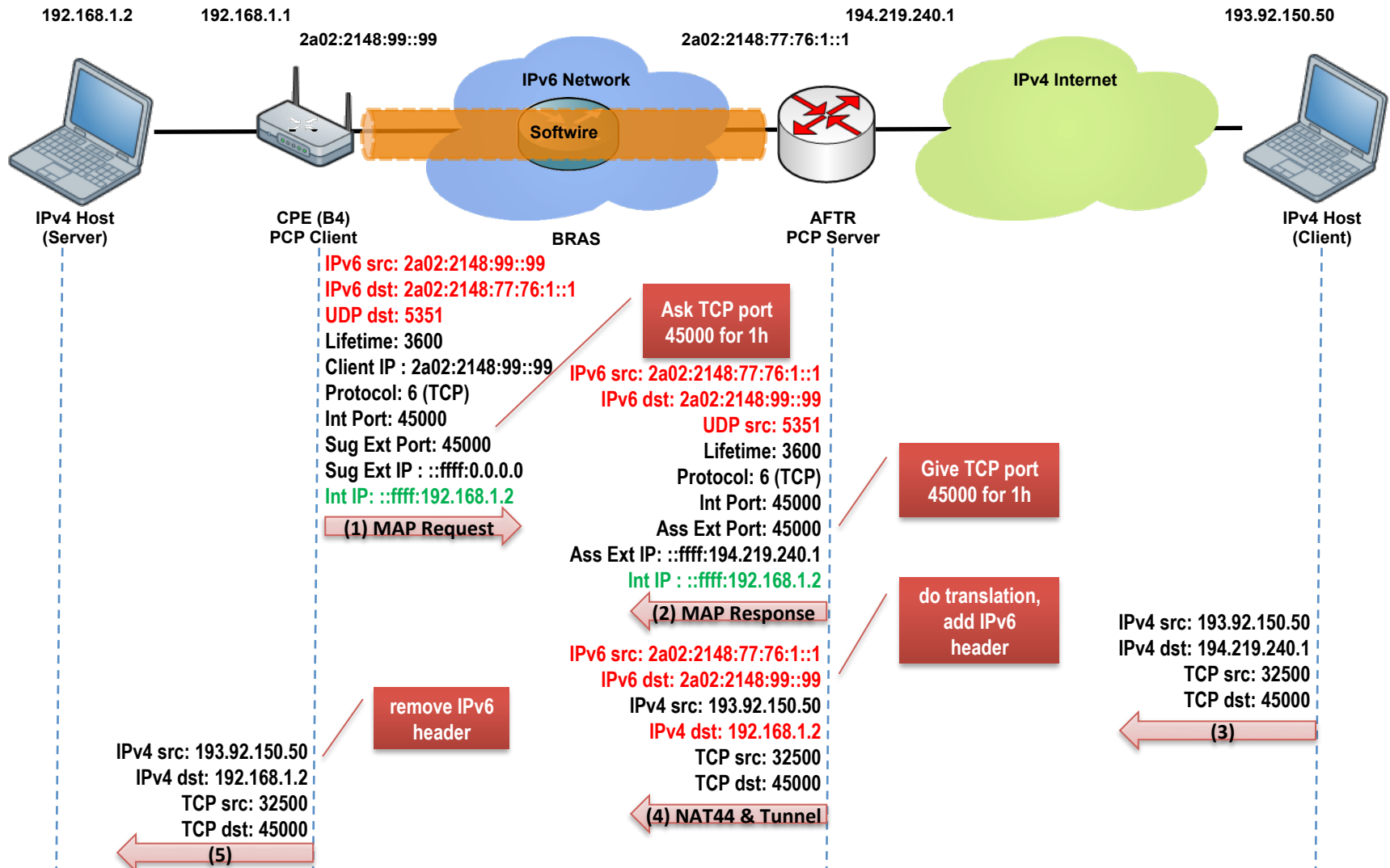
PCP might become an all-around protocol for signaling (like BGP)

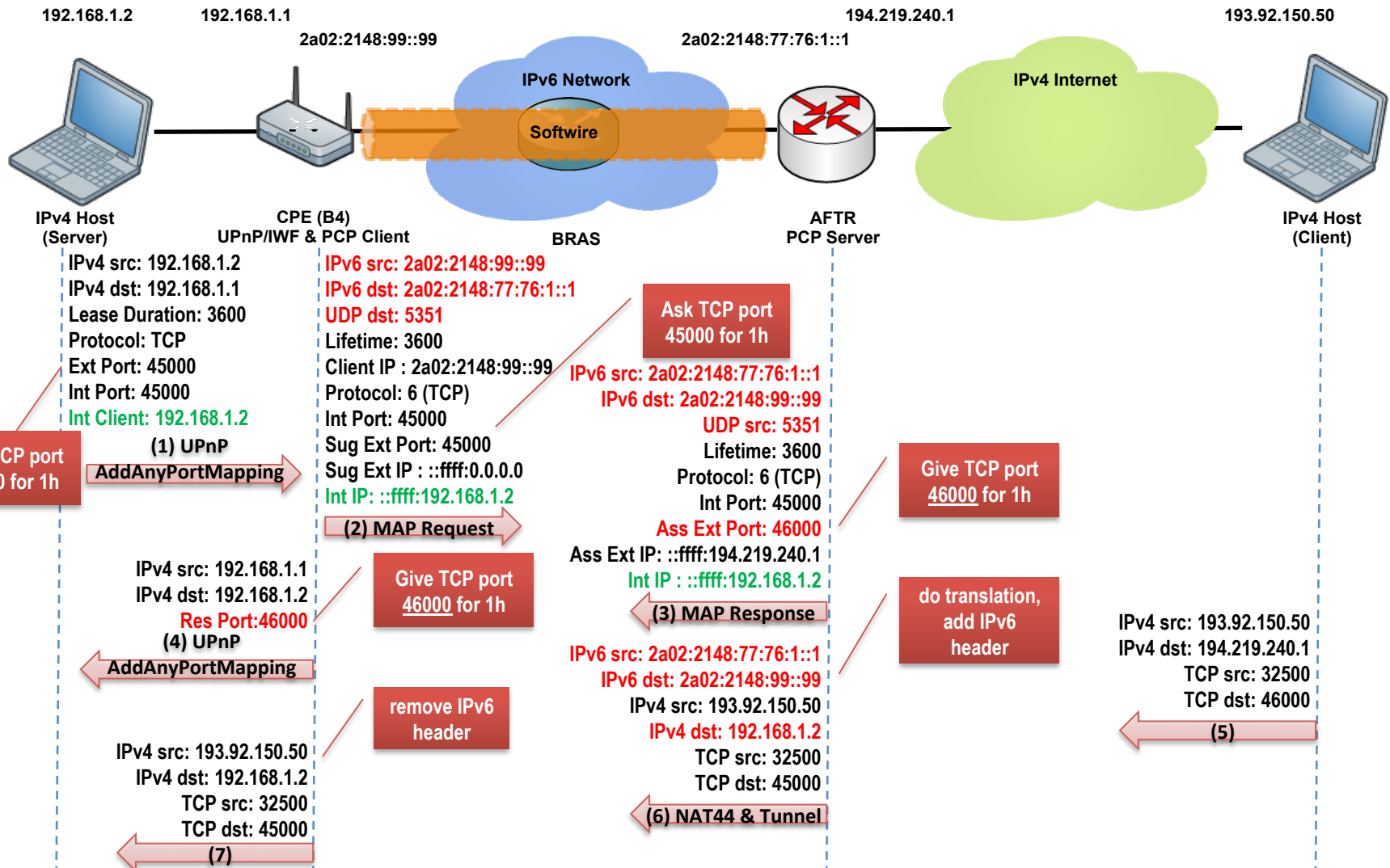
PCP main objective

- ❑ Enable subscriber applications to receive incoming connections in coordination with an ISP NAT/Firewall device
- ❑ Enable an explicit communication between CPE/applications and the ISP NAT/Firewall device, instead of using NAT traversal techniques (STUN/TURN/ICE)
- ❑ Good potential to become the “carrier-grade” evolution of UPnP and NAT-PMP

PCP & DS-Lite

- ❑ In plain mode, B4 element acts as PCP-client or PCP-proxy for IPv4 clients behind it and will send requests for PCP mappings using the THIRD_PARTY option





Common Request Header					
Version:	2	R:	0	Opcode: 1	Reserved
Requested Lifetime:		3600			
PCP Client's IP address:		2a02:2149:ff02:ca4b:f520:adfa:894:1b99			
<input type="radio"/> ANNOUNCE <input checked="" type="radio"/> MAP <input type="radio"/> PEER Opcode Request					
Protocol:	6	Reserved			
Internal Port:	50007	Suggested External Port:	50007		
Suggested External IP address:		::ffff:0.0.0.0			
Option(s)					
Option Code:	1	Reserved	Option Length:	16	
Data:	::ffff:192.168.1.2				
<input type="button" value="Add an Option"/> <input type="button" value="Delete"/> <input type="checkbox"/> Filter option?					

A custom page was created on the CPE in order to start experimenting with PCP

PCP

Tests

Create a NAT port-forwarding entry for a web server on internal port 50005

Create a NAT port-forwarding entry for a web server on internal port 50006

Protocol	Internal IP Address	Internet Port	External IP Address	External Port	Lifetime	Created by
TCP	192.168.1.2	50005	194.219.240.252	50005	Infinite	User
TCP	192.168.1.2	50006	194.219.240.252	34873	Infinite	User



Implementation issues

- ☐ Many PCP servers seem to be stuck at draft versions (RFC out from April)
- ☐ Support mostly for MAP opcode
- ☐ Main focus on PCP client and PCP/UPnP-IGD IWF functionality
- ☐ PCP proxy not preferred (any apps out there?)
- ☐ PREFER_FAILURE by default on
- ☐ Many bugs with lifetimes and timeouts

Still a lot to be done in PCP

Nevertheless,
DS-Lite and PCP work!

```
::6:Jun:2008 prepare  
::6:Jun:2010 sign in to 2a02:2148:100:a01::22  
::6:Jun:2011 take flight lessons  
::6:Jun:2012 we have lift-off  
::1:Apr:2013 on fire - it's not a joke  
::6:Jun:2013 reach the troposphere  
::1:Apr:2014 burn out - it's still not a joke  
::6:Jun:2014 above the stratosphere  
::6:Jun:2015 where no man has gone before
```

What are **your** plans?

- ☐ IPv6 internal tests
- ☐ /29 from RIPE & ipv6.forthnet.gr
- ☐ IPv6 pilot (W6D)
- ☐ IPv6 production (W6L)
- ☐ IPv6 production & DS-Lite pilot
- ☐ 10% of subscribers use IPv6
- ☐ DS-Lite & PCP production
- ☐ 30% of subscribers use IPv6
- ☐ ?

Thank you!

Q & A