

Dual-Stack lite

Alain Durand

May 28th, 2009

Part I: Dealing with reality

A dual-prong strategy

IPv4 reality check: completion of allocation is real



After completion:

**Existing IPv4 addresses will not stop working.
Current networks will still operate.**

IPv6 reality check: the IPv4 long tail

- Post IPv4 allocation completion:
 - Many hosts in the home (eg Win 95/98/2000/XP, Playstations, consumer electronic devices) are IPv4-only.
 - They will not function in an IPv6-only environment.
 - Few of those hosts can and will upgrade to IPv6.
 - Content servers (web, email,...) hosted on the Internet by many different parties will take time to upgrade to support IPv6.

Dealing with both realities: a two prong approach

① Embrace IPv6

- Move as many devices/services to IPv6 as possible to lower dependency on IPv4 addresses













② Build an IPv6 transition bridge for the IPv4 long tail

- Goal:
 - Provide IPv4 service without providing a dedicated IPv4 address
- Technology:
 - Leverage IPv6 access infrastructure
 - Provide only IPv6 addresses to endpoint
 - Share IPv4 addresses in the access networks
 - DS-lite: IPv4/IPv6 tunnel + provider NAT

Part II: Plan A, B, C, ...

Lessons learned

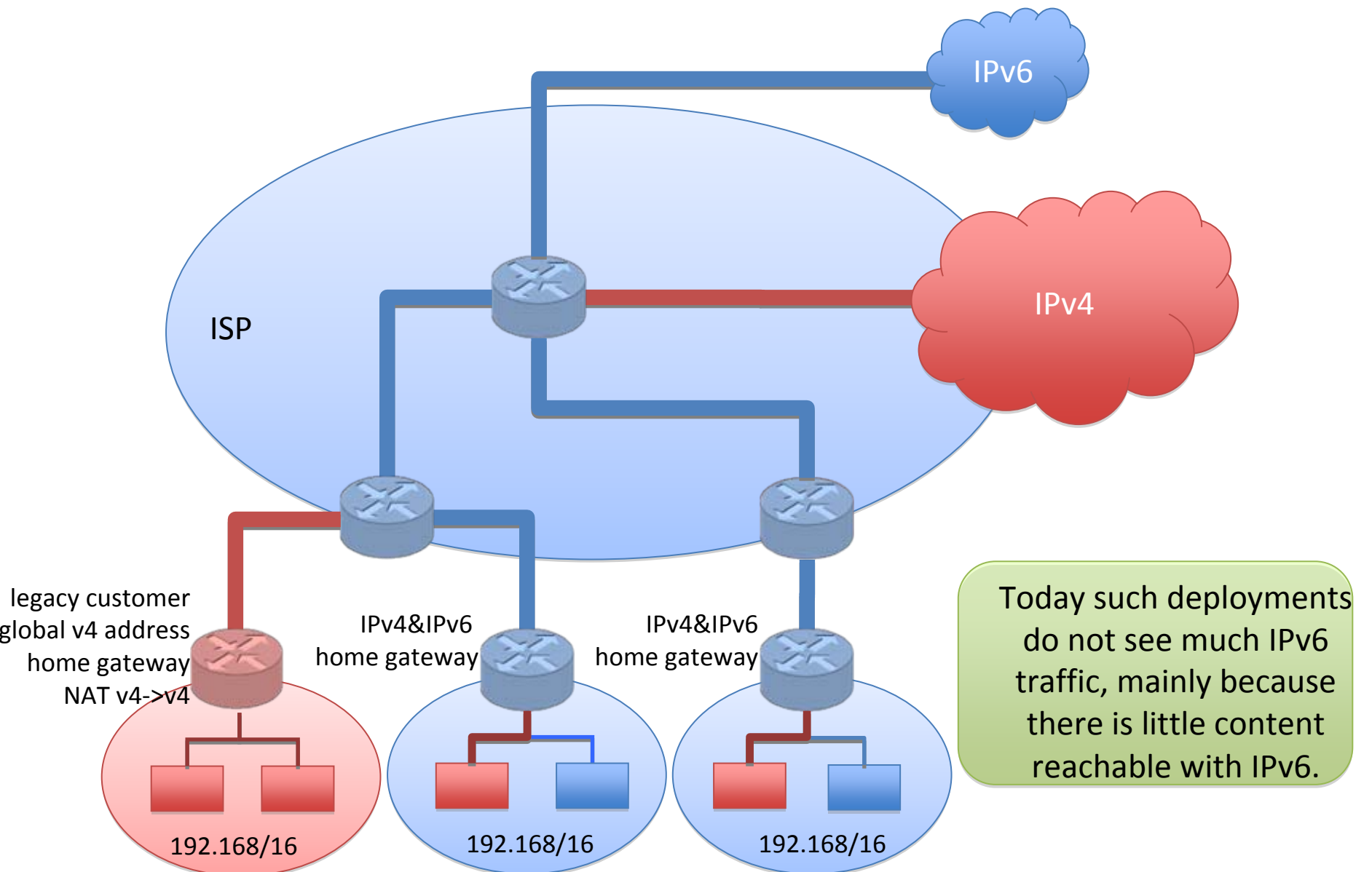
Provisioning color code

	IPv4-only	dual stack provisioned	dual stack*, IPv6-only provisioned
device			
link			
router			
network			

* devices with pure IPv6-only code are out of scope

Plan zero: dual-stack

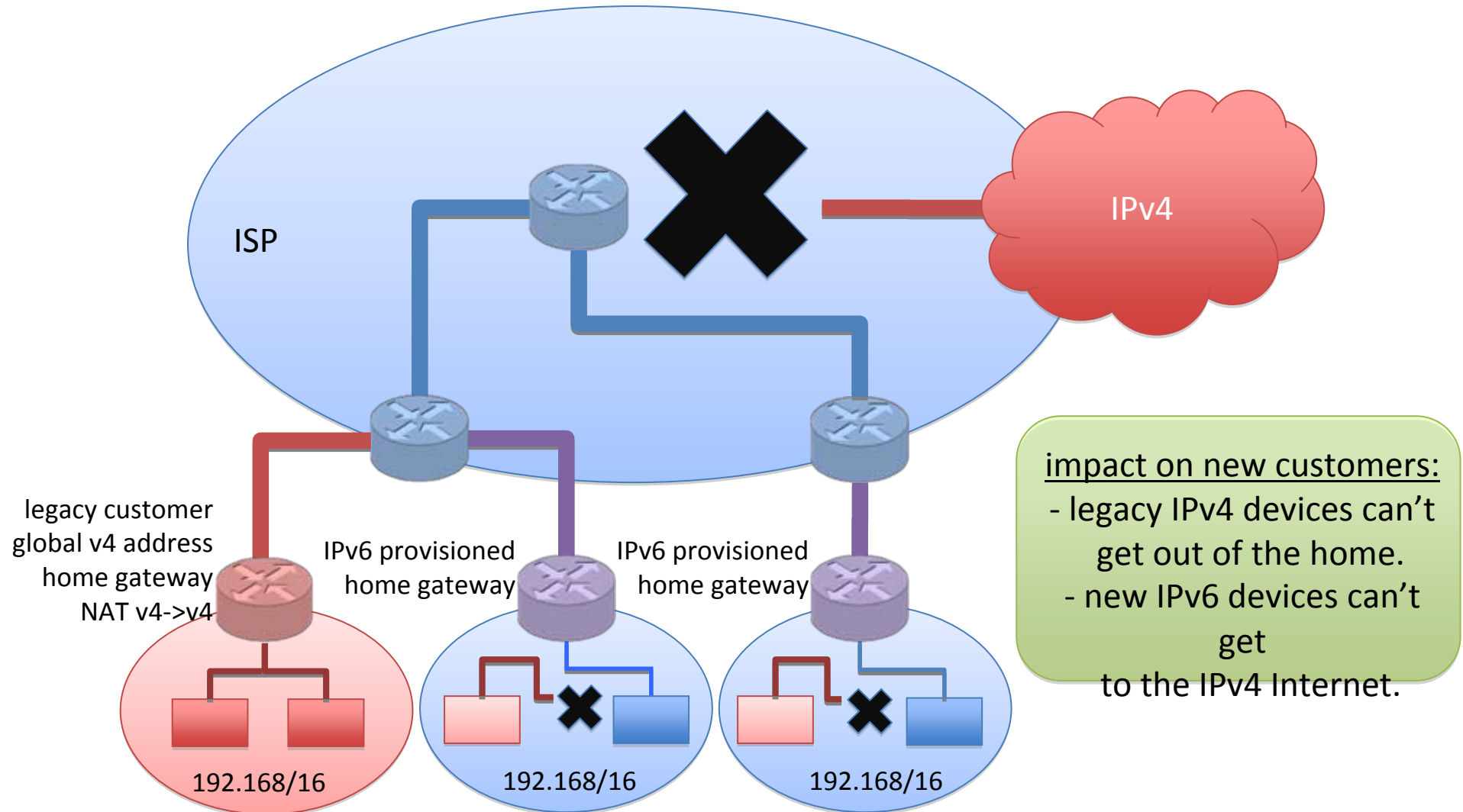
After IPv4 IANA completion, there will not be enough IPv4 addresses to sustain this model.



Plan A: dual-stack core

new customers are provisioned
with IPv6-only but no IPv4 support

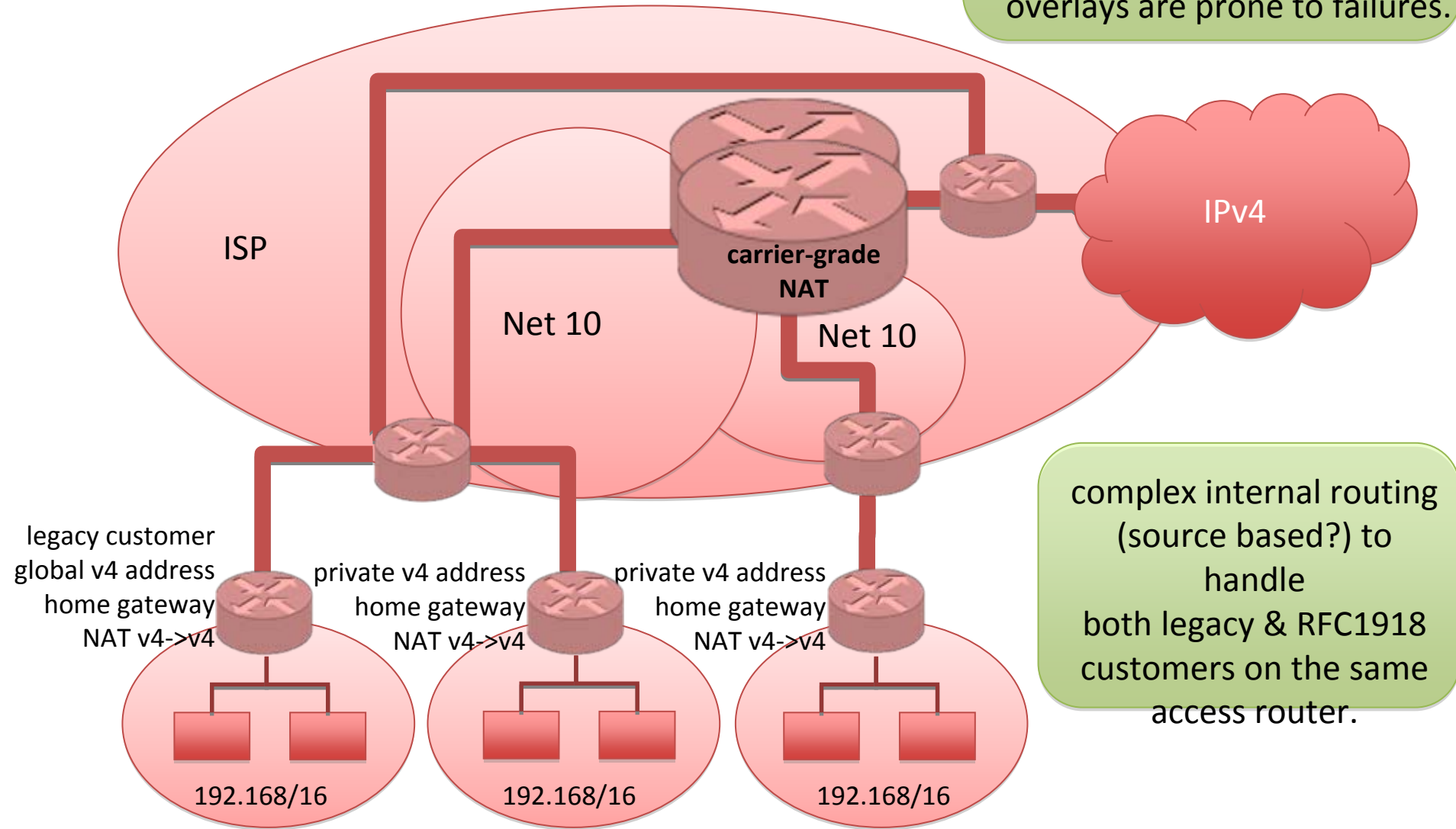
lots of broken paths...



Plan B: double NAT

new customers are provisioned
with overlays of RFC1918

- two layers of NAT
- no evolution to IPv6
- network gets increasingly complex to operate.
- Intersections of Net 10 overlays are prone to failures.

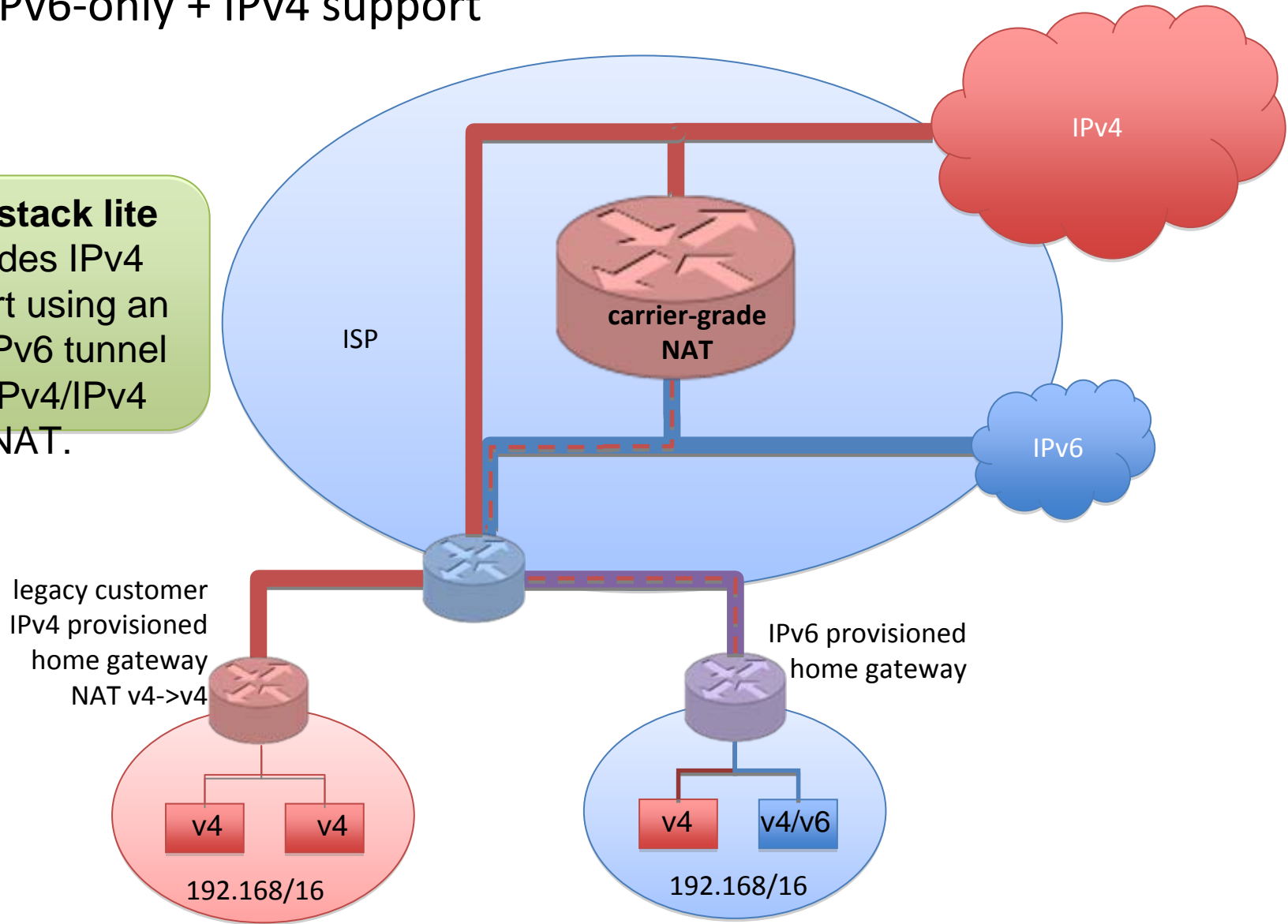


Plan C: dual-stack lite

new customers can be provisioned with IPv6-only + IPv4 support

- simplifies network operation
- provides an upgrade path to IPv6

Dual-stack lite provides IPv4 support using an IPv4/IPv6 tunnel to a IPv4/IPv4 NAT.

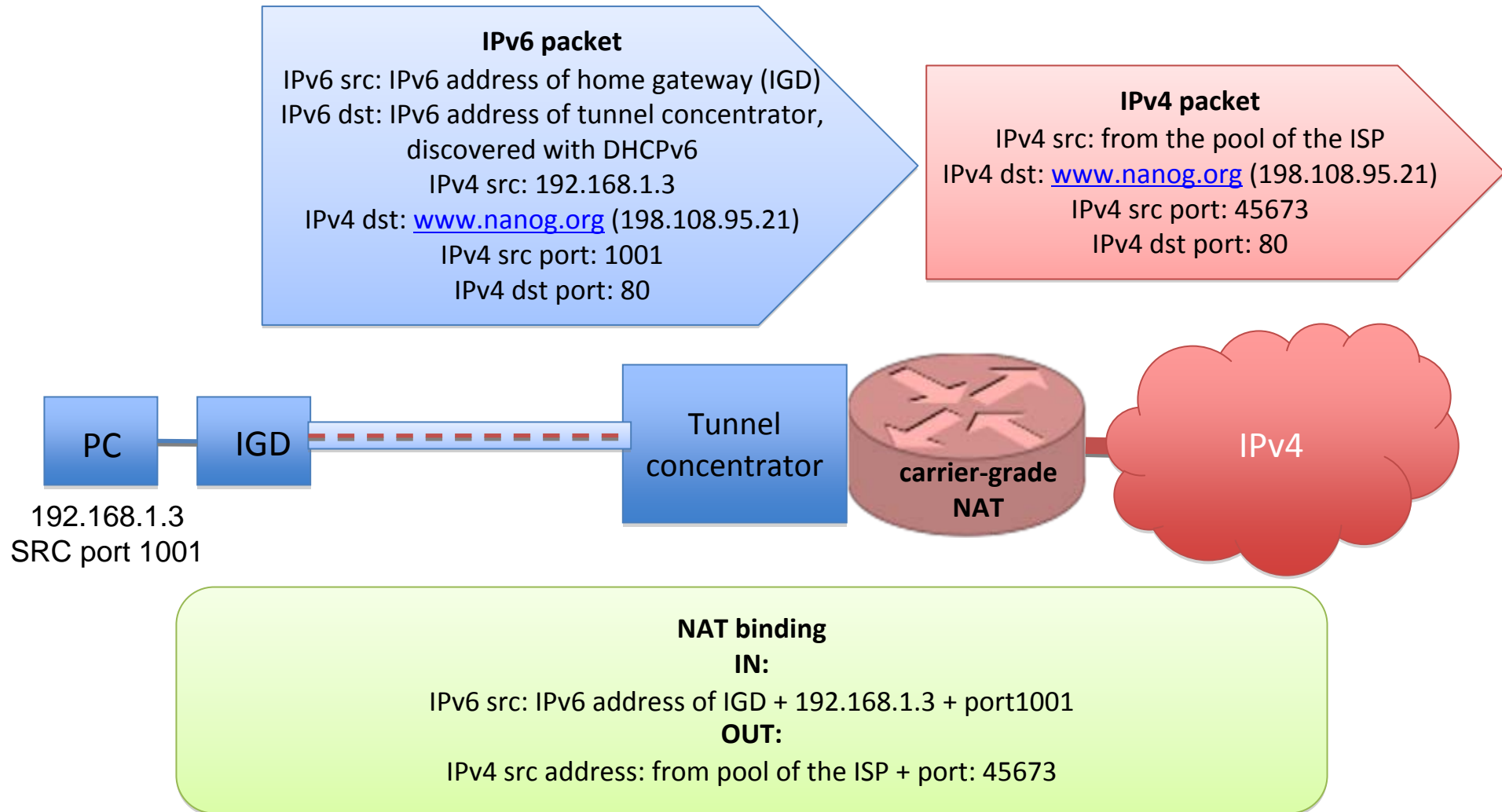


Part III: DS-lite technology

Combining two well-known
technologies:
NAT + Tunneling

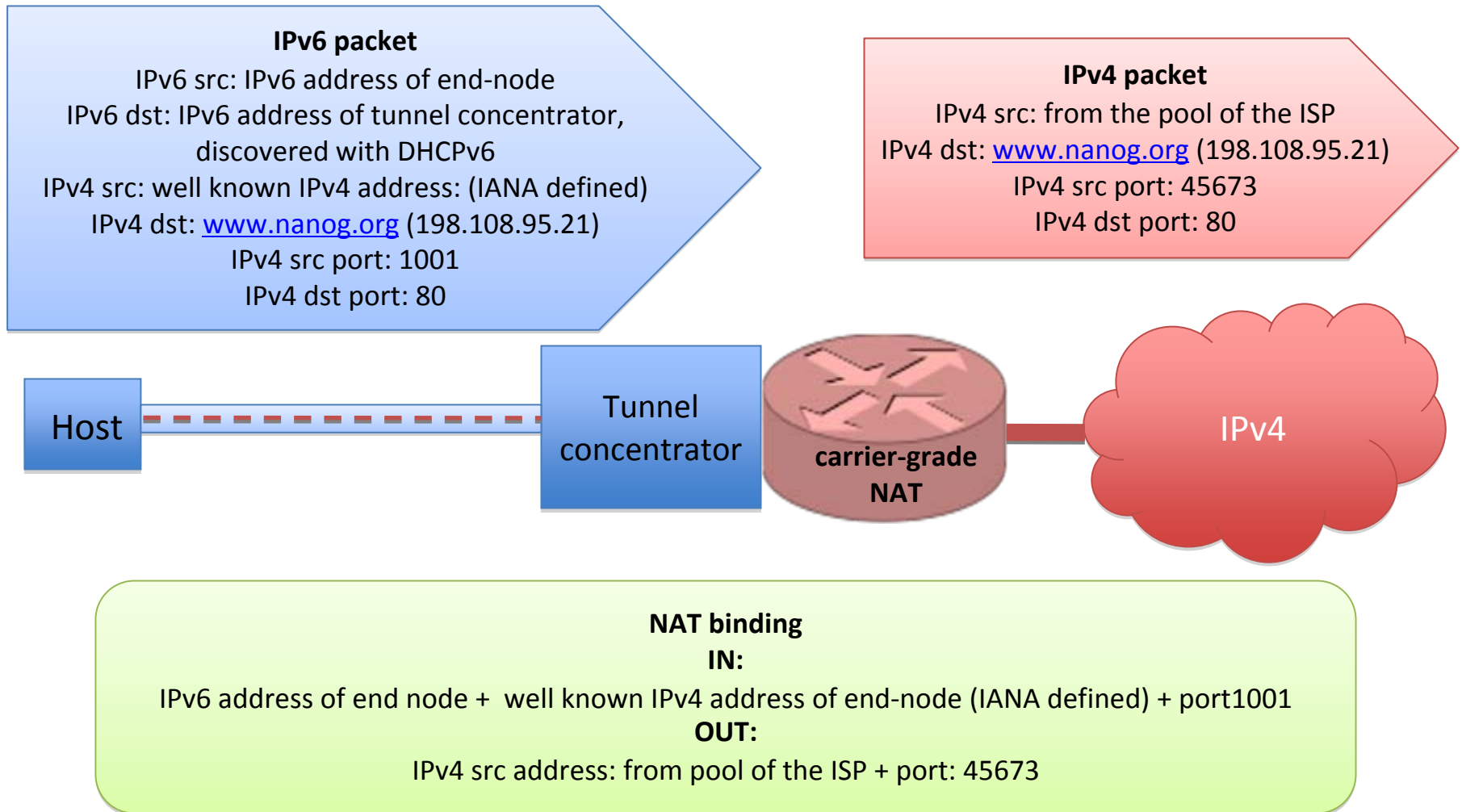
Gateway-based scenario:

IGD are provisioned with IPv6-only + IPv4 support for the home PC from a carrier-grade NAT



End-node scenario:

Dual-stack capable end-nodes are provisioned with IPv6-only + IPv4 support from a carrier-grade NAT



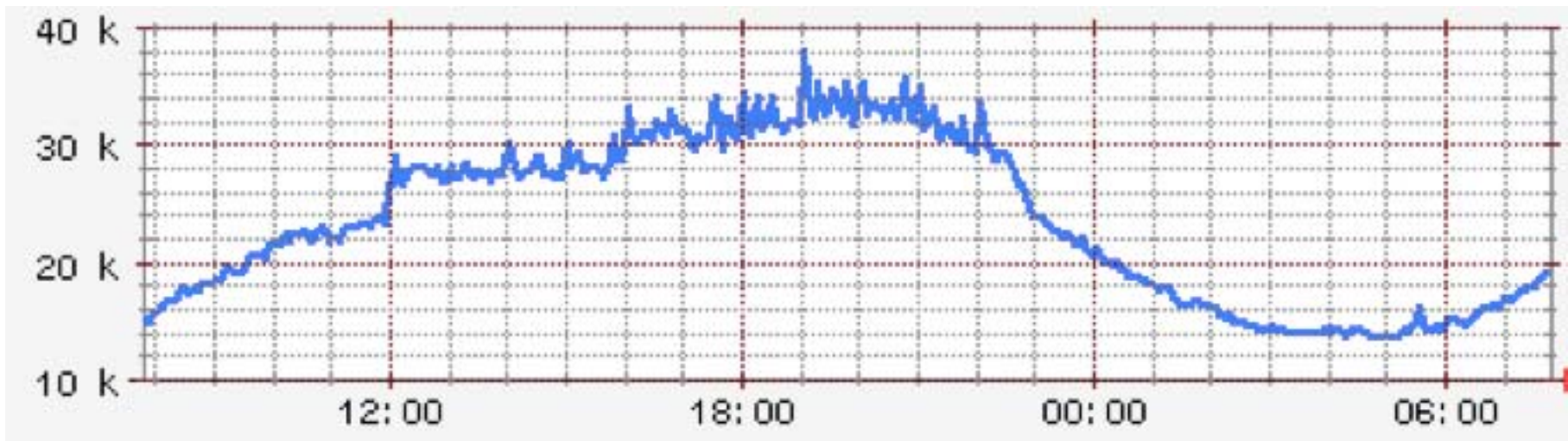
Part IV: TCP/UDP port consumption

Measurements

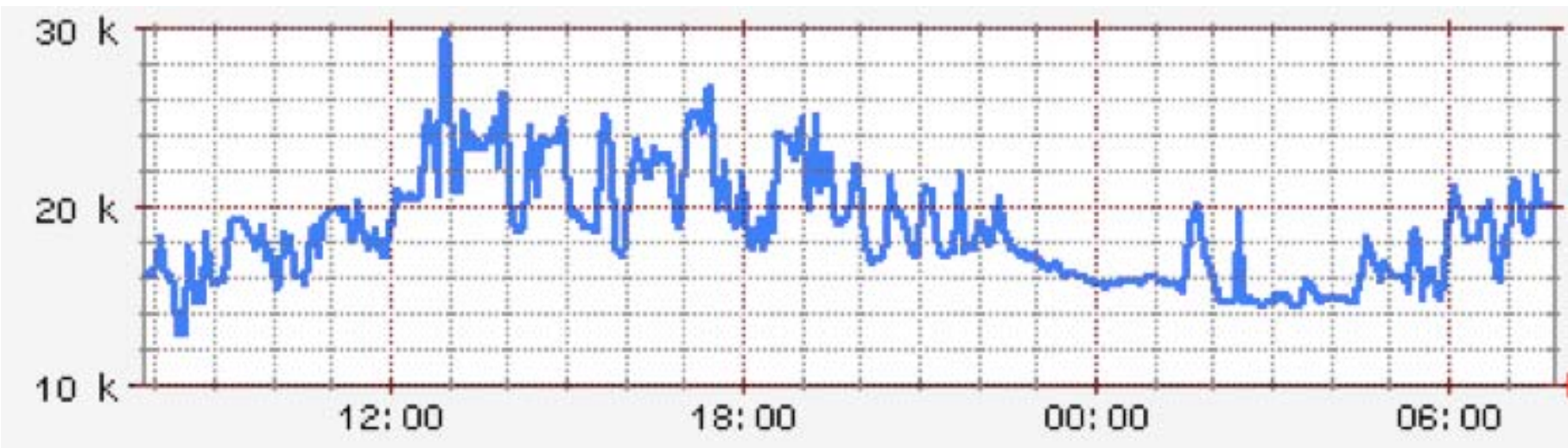
Measurements

- Measurement campaign performed on 2008-11-10
- Data collected behind a CMTS with 8000 subscribers
- Caveats:
 - Data was collected on only one point in the network
 - Measurement methodology still needs to be tuned
 - Results need to be compared to other studies
 - “Your mileage may vary”

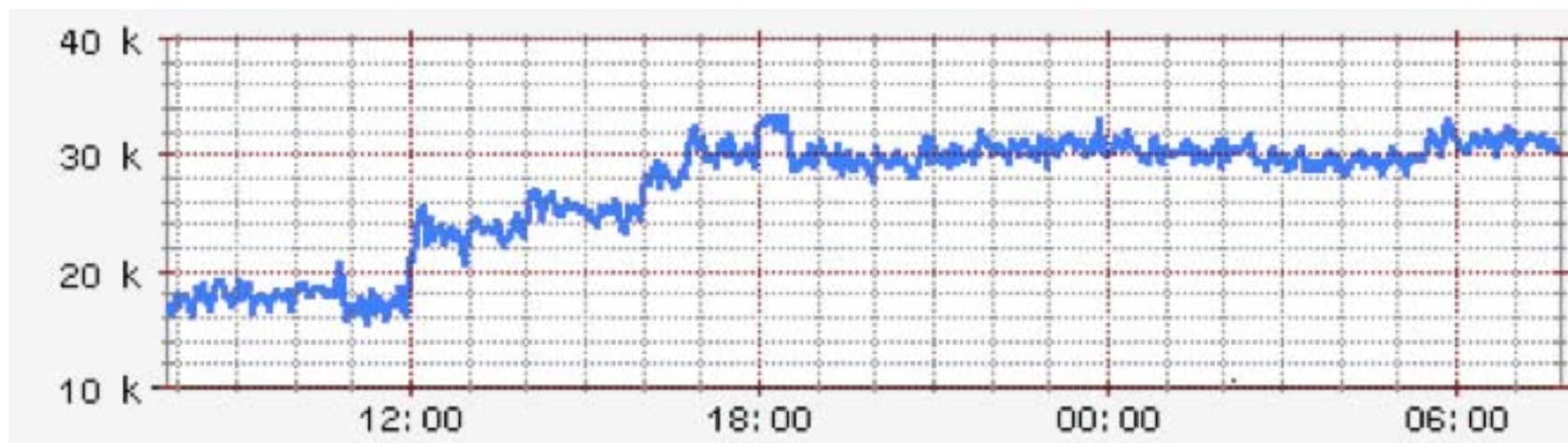
TCP 'outgoing' ports statistics on a 8000 subscriber sample



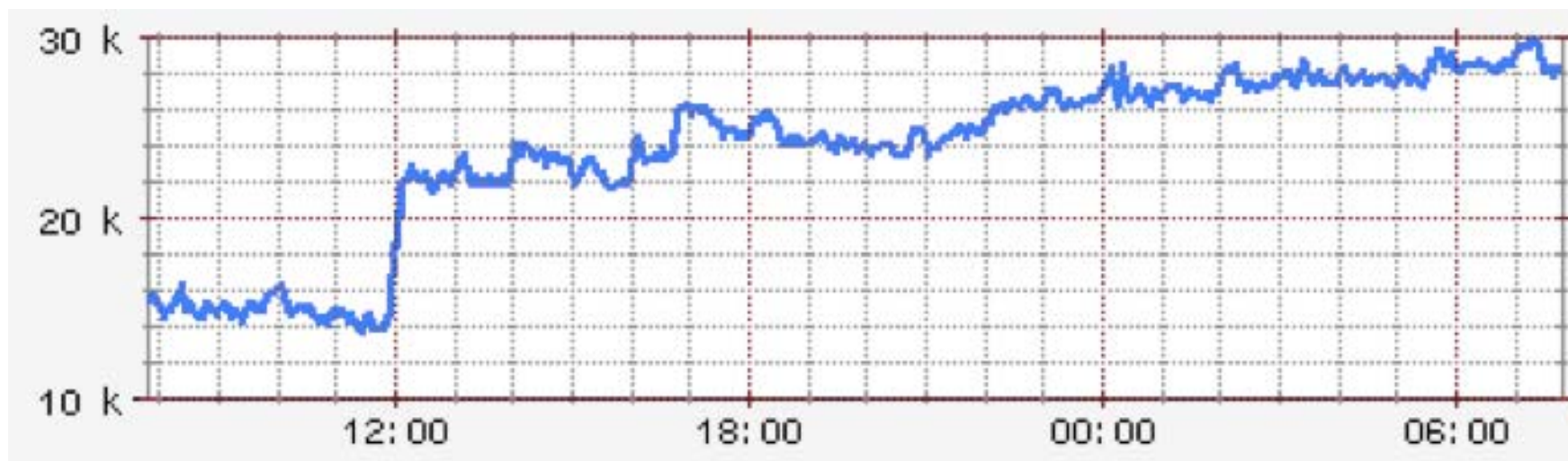
TCP 'incoming' ports statistics on a 8000 subscriber sample



UDP 'outgoing' ports statistics on a 8000 subscriber sample



UDP 'incoming' ports statistics on a 8000 subscriber sample



Analysis

- Maximum average: 40,000 ports/protocol/direction
- It translates into a maximum of 5 ports consumed per user on **average** in each direction for both TCP & UDP.
Total: 20 ports/user on average
- This needs to be compared with the hundreds/thousands of ports that can be consumed at **peak** by a single user browsing a Web 2.0/AJAX site.
- One needs to keep this analysis in mind when designing a port distribution mechanism for a carrier-grade NAT.

Part V: DS-lite standardization

Draft-ietf-softwire-dual-stack-lite-00.txt

DS-lite Status

- IETF

- Latest draft:
 - `draft-ietf-softwire-dual-stack-lite-00.txt`
- IETF softwire WG has been re-chartered to standardize DS-lite.

- Implementations

- IGD: Open source code (Open-WRT) for a Linksys home router
- CGN: Vendor code, open source project started

IPv4 port distribution

- Measurements:
 - Average #ports/customer < 10 (per transport protocol)
 - Peak #ports/customer > 100? > 1000? > 5000?
- Do not dimension for peaks, but for average!
 - No cookie cutter approach
 - Large dynamic pool of ports shared by many customers
- Customers want to choose their own applications
 - CGN MUST not interfere with applications, eg avoid ALGs,...
 - Need to support incoming connections
 - Small static pool of reserved ports under the control of customers

Port forwarding & A+P extensions

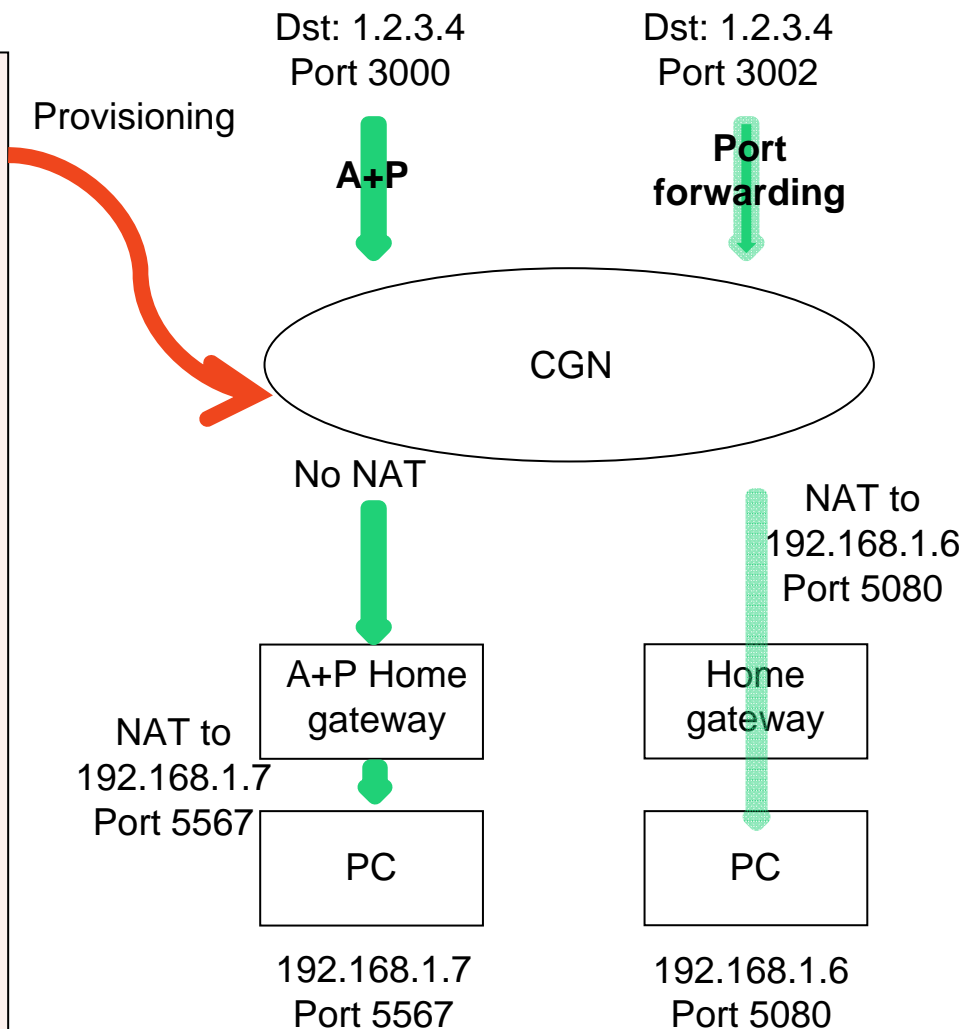
ISP portal

Address & port control tab

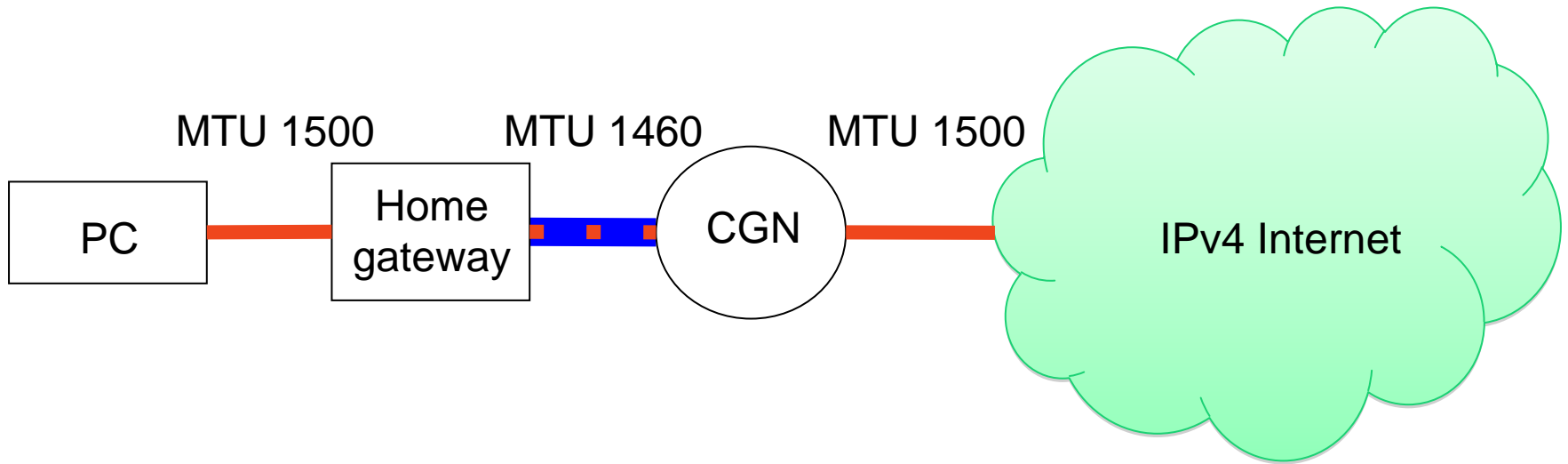
User: X

External IPv4 address: 1.2.3.4

Port	A+P	Port forwarding	Internal IP	Port
3000	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3001	<input type="checkbox"/>	<input checked="" type="checkbox"/>	192.168.1.5	80
3002	<input type="checkbox"/>	<input checked="" type="checkbox"/>	192.168.1.6	5080
3003	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
3004	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
...				



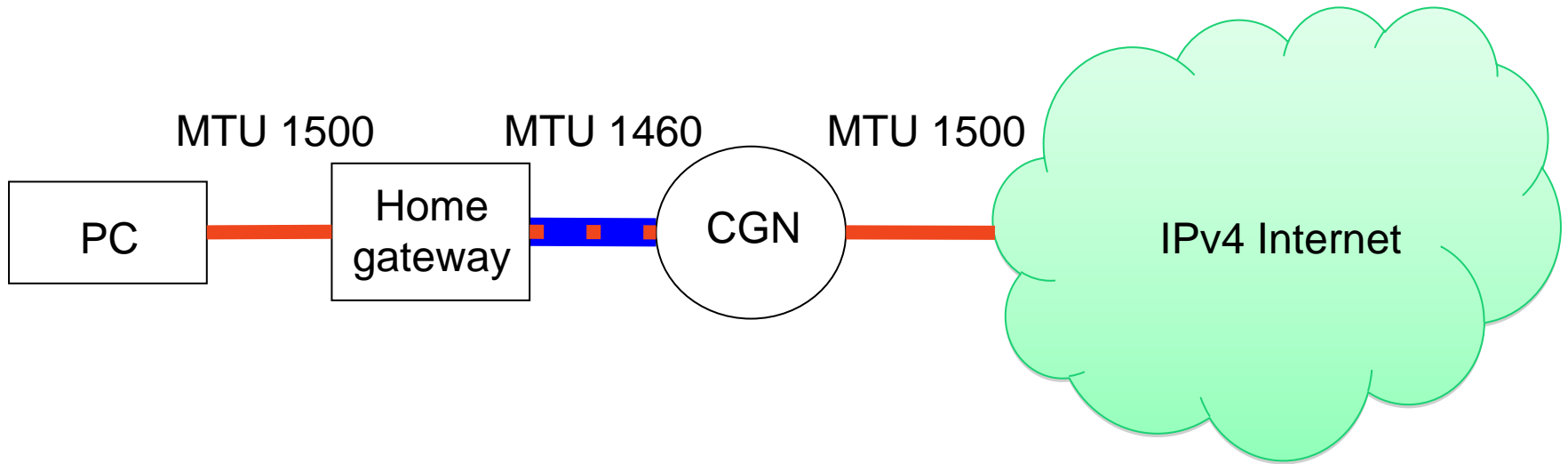
Issues with MTU



pMTU discovery does NOT work over the tunnel

IPv4 fragmentation needs to be avoided

Dealing with MTU



For TCP: CGN rewrites the TCP MSS in the SYN packet

For UDP: HGW & CGN do IPv6 fragmentation/reassembly over the tunnel

Part VI: Generic issues with CGN

Applies to DS-lite, NAT444, NAT64,IVI,...

UPnP

- Typical UPnP application will:
 - Decide to run on port X
 - Ask IGD to forward port X traffic
 - If IGD declines, try again with X+1
 - After 10 or so attempts, abort
- This will NOT work with any IPv4 address sharing mechanism (NAT444, DS-lite, NAT64, IVI, A+P,...)
- NAT-PMP has a better semantic: IGD can redirect the application to use an alternate available port
- UPnP forum is reported to be addressing this issue

Security issues relative to CGN

- Port number information is necessary for full identification
 - Need to log port numbers on the receiving side
 - Need to log NAT bindings on CGN
- CGN needs to enforce per customer limits either on new connection rate or maximum number of sessions
- User authentication on service provider CGN may not be necessary, users get authenticated at the IPv6 access layer. A simple ACL on the CGN to limit access to the service provider customers seems to be sufficient. 3rd party CGNs may have different requirements.
- HGW & CGN need to enforce that customer IPv4 addresses inside of IPv6 tunnel are indeed RFC1918 addresses

Other security issues

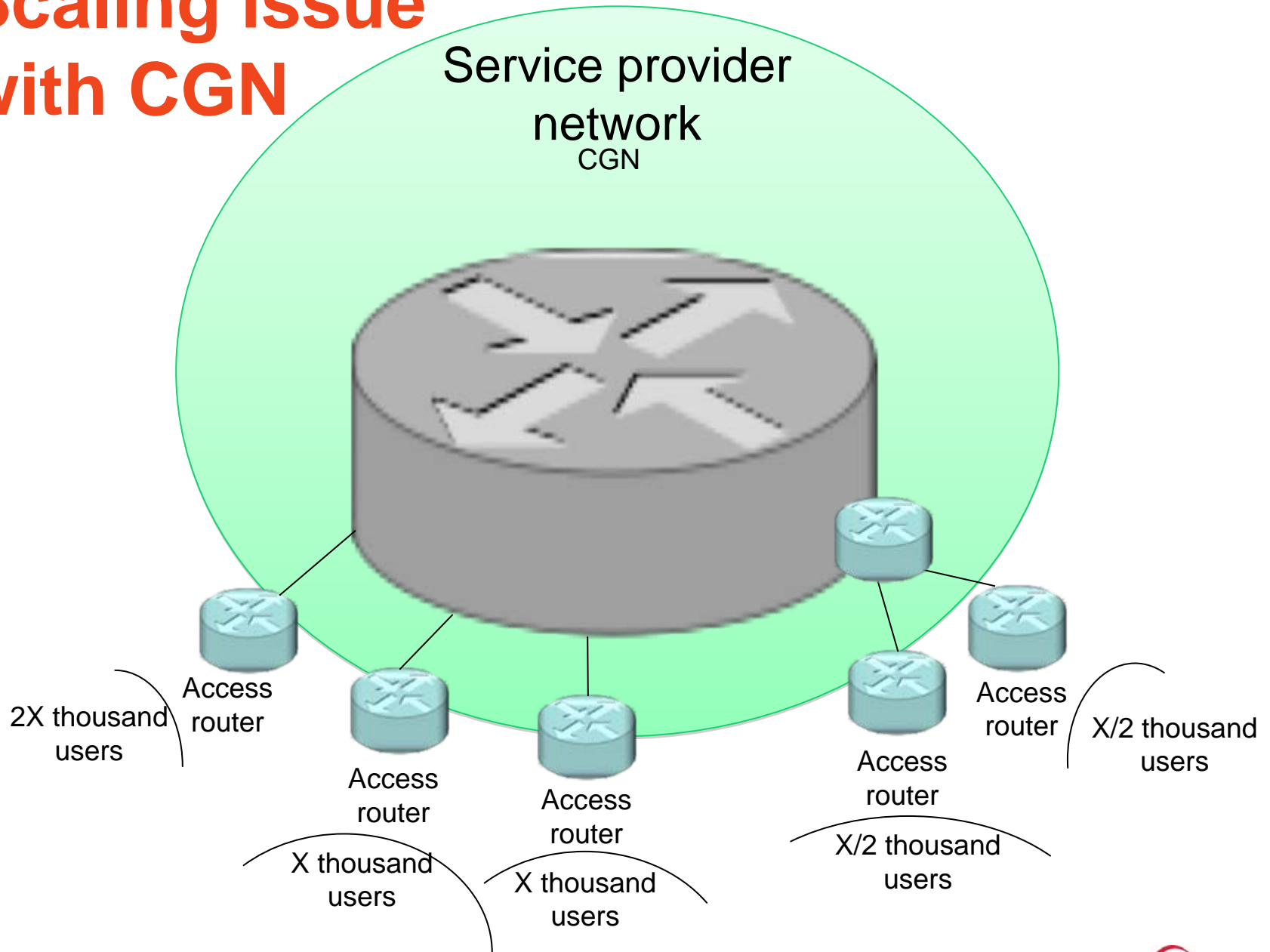
- The Internet community needs to deal with Web sites that put IPv4 address in penalty box after a number of unsuccessful login attempts.
- More generally, the community need to revisit notion that an IPv4 address uniquely identifies a customer.

Part VII

DS-lite deployment model

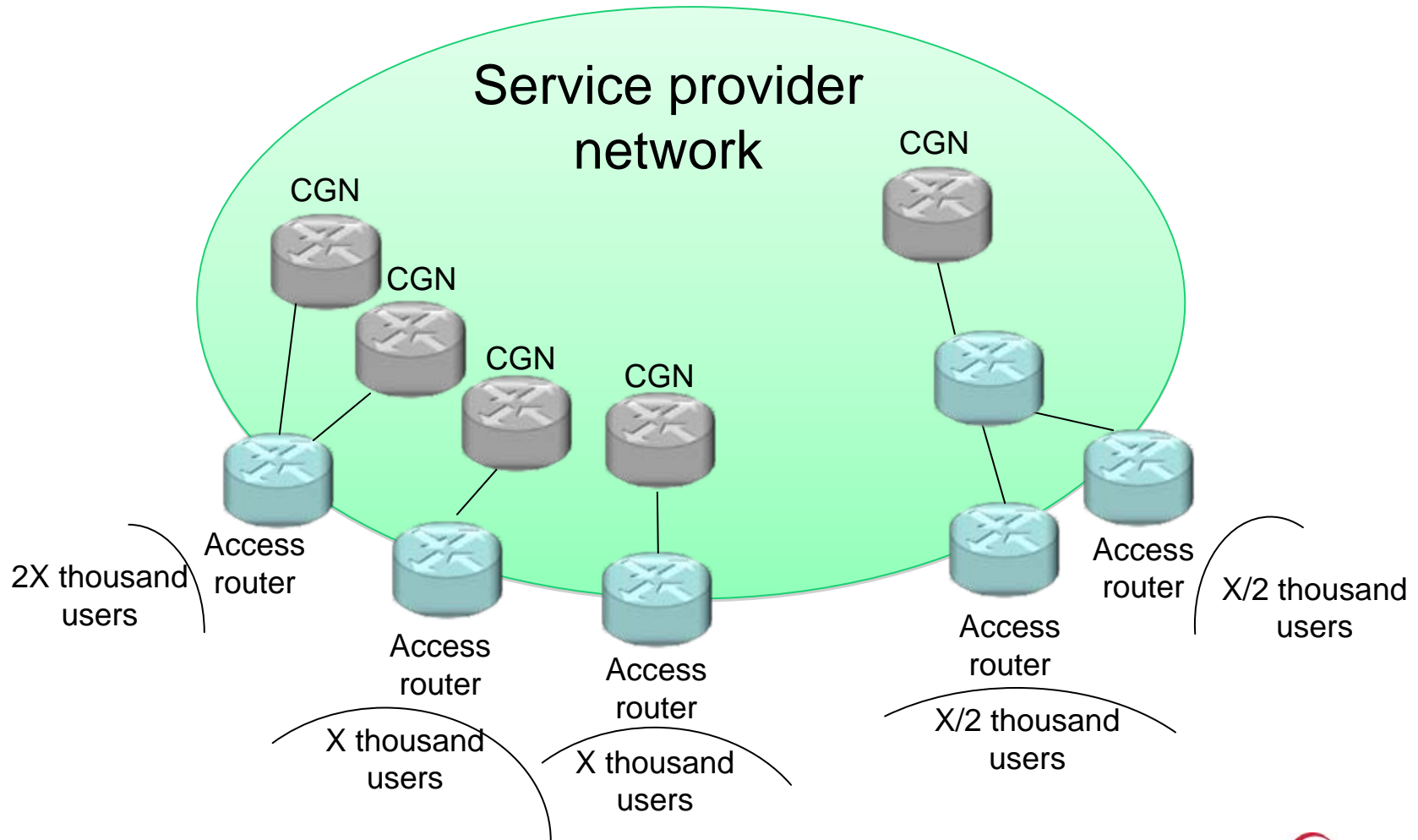
Horizontal Scaling

Scaling issue with CGN



Horizontal scaling with DS-lite

- DHCPv6 option to configure tunnel end-point
- Enable sending the traffic to as many CGNs as necessary



Part VIII

DS-lite demo at LACNIC XII

Thanks to:
Yiu Lee, Carl Williams, Anthony Veiga
ISC
LACNIC, Roque Gagliano

LACNIC/Comcast DS-lite demo

LACNIC

Lacnic
v4/v6
Router

Lacnic dual-stack wired network

IPv6 static route to /56 IPv6 prefix

1 IPv6 address
/56 IPv6 prefix

1 IPv4 address

DHCPv6:

- IPv6 address of home GW
- /64 DHCPv6 prefix delegation
- DNS server IPv6 address
- CGN IPv6 address

IPv6 router
DNS/DHCPv6
server

DS-lite
CGN

/32 IPv4
addresses for
NAT pool

/64 DHCPv6-PD

Home GW

Wifi
SSID
Lacnicxii-dslite

PC
1

PC
20

/64 IPv6
prefix

Color code

IPv6

Dual-
stack

IPv4

COMCAST

Lacnic IPv6 network

