A+P Lite

How to Keep CGNs from Breaking the Internet

APRICOT / Feb 24 2009

Randy Bush <randy@psg.com>

Alain Durand <alain_durand@cable.comcast.com>

Olaf Maennel <olaf@maennel.net>

Problem Statement

Broadband providers will not have enough IPv4 space to give one IPv4 address to each CPE or terminal so that every consumer has usable IPv4 connectivity.

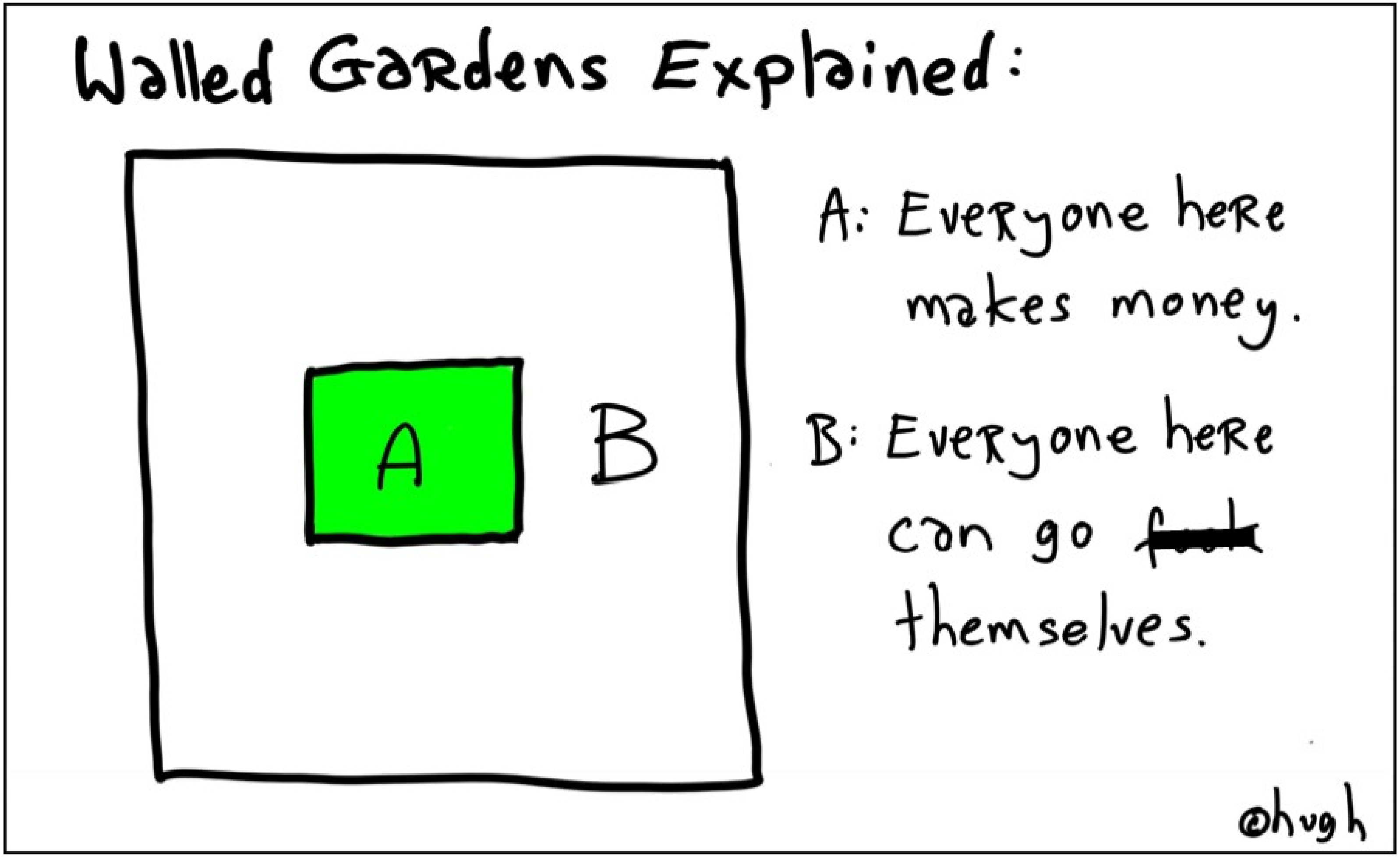
Large-Scale NAT (LSN)

- LSN (formerly CGN) are NATs in the core of the provider's network
- NATs did not scale to Carrier Grade, no big surprises
- Customer Premisis Equipment (CPE) has 4to6 NAT and the core re-NATs 6to4 for v4 destinations

LSN Breaks the Net

- This cause problems for the carrier, but also for the whole internet, as these captive customers can not use new protocols
- NAT in middle of net has all of the problems of a smart core, the Telco model
- Walled gardens here we go!

I Googled “Walled Garden”



Walled Gardens Explained:

A: Everyone here makes money.

B: Everyone here can go [redacted] themselves.

Captive Users

- This is the business model of User as Consumer
- Internet becomes Television
- Media is Controlled (DRM)
- Protocol innovation Stops
- RFC 1918 is totally deployed
- The Internet of the Telcos

This Is Not Inevitable

Keep the Power of Choice in the Hands of the Users!

Allow the NAT to be "flexible"

A+P in One Slide

- Goal: mechanism required that **customer can control their "fate".**

- "Steal" bits from Ports and use it for addressing. Same as LSN.
- But do it at the User CPE!
- Thus, extend end-to-end connectivity (at least for some ports) to end-user!

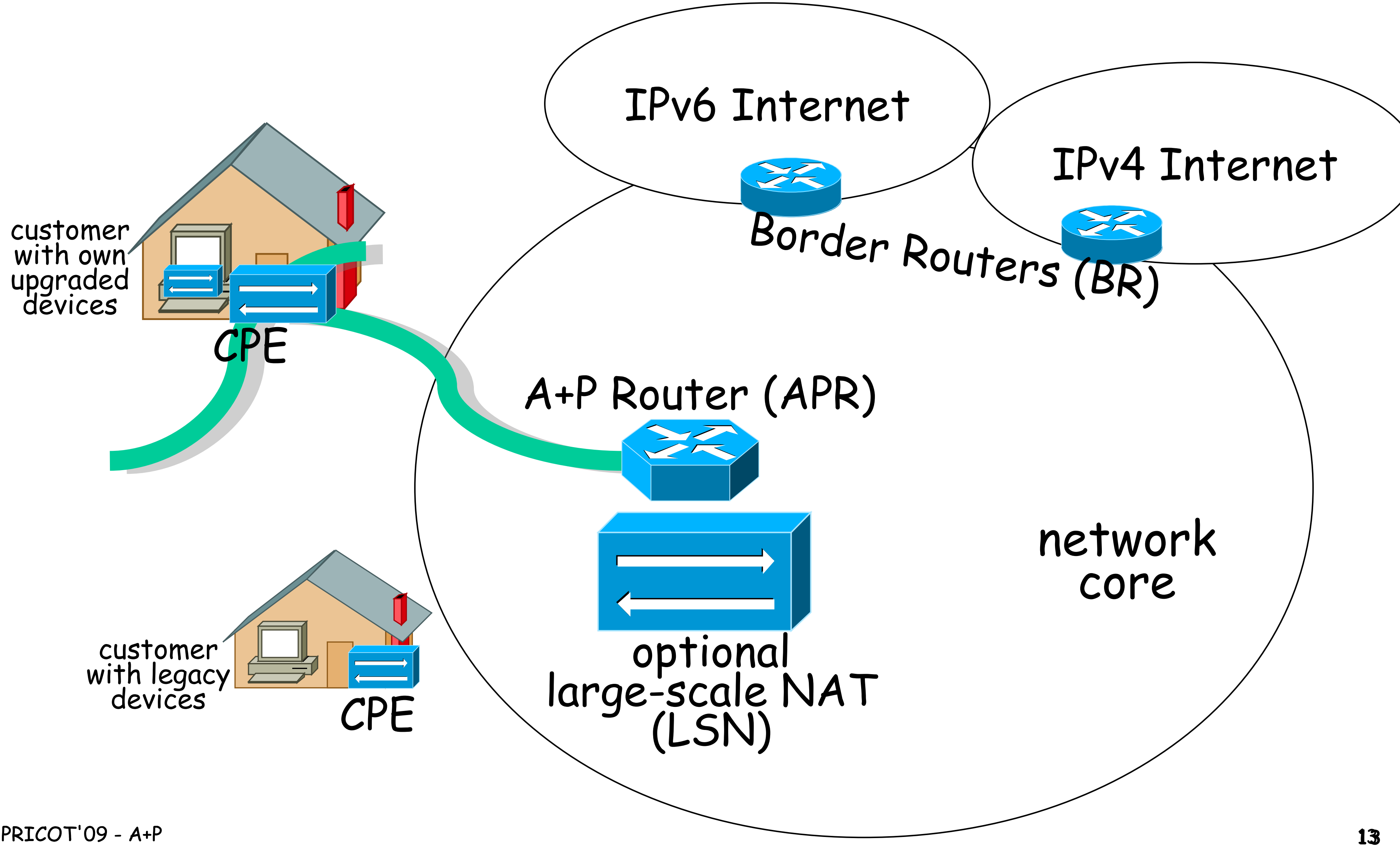
Same Port-Count Issues as LSN

- Trade-off between port efficiency and signaling
- Measurement studies show port-use per residential customer ~100, peaking at ~700
- We are out of addresses, so we share and this is the consequence. No magic

Separate the Functions

- **Encaps / Decaps**
  - "Softwire" (transport pkts from/to CPE)
  - End-user has control over some *untranslated* ports end-to-end
- **NAT**
  - Inevitable to connect legacy devices
  - But: flexible of where NATing is done

A+P Lite Terminology


IPv6 Internet
IPv4 Internet
Border Routers (BR)
customer with own upgraded devices
CPE
A+P Router (APR)
network core
optional large-scale NAT (LSN)
customer with legacy devices
CPE

Alain (Comcast) Says

It is expected that the home gateway is either software upgradable, replaceable or provided by the service provider as part of a new contract.

Does Not Have to be True for All Providers

Constraints for Possible Solutions

Constraints (I)

1) **End-to-end is under customer control.** Customers shall have the possibility to send/receive packets unmodified and deploy new application protocols at will. IPv4 address exhaustion is no clearance to break the Internet's end-to-end paradigm.

2) **End-to-end transparency through multiple intermediate devices.** Multiple gateways should be able to operate in sequence along one data path without interfering with each other.

3) **Incremental deployability and backward compatibility.** The approaches shall be transparent to unaware users. Devices or existing applications shall be able to work without modification. Emergence of new applications shall not be limited.

4) **Automatic configuration/administration.** There should be no need for customers to call the ISP and tell them that they are operating their own A+P-gateway devices. Customers/mobile phone users are NOT supposed to lookup assigned ports manually on websites and then configure them on devices or applications.

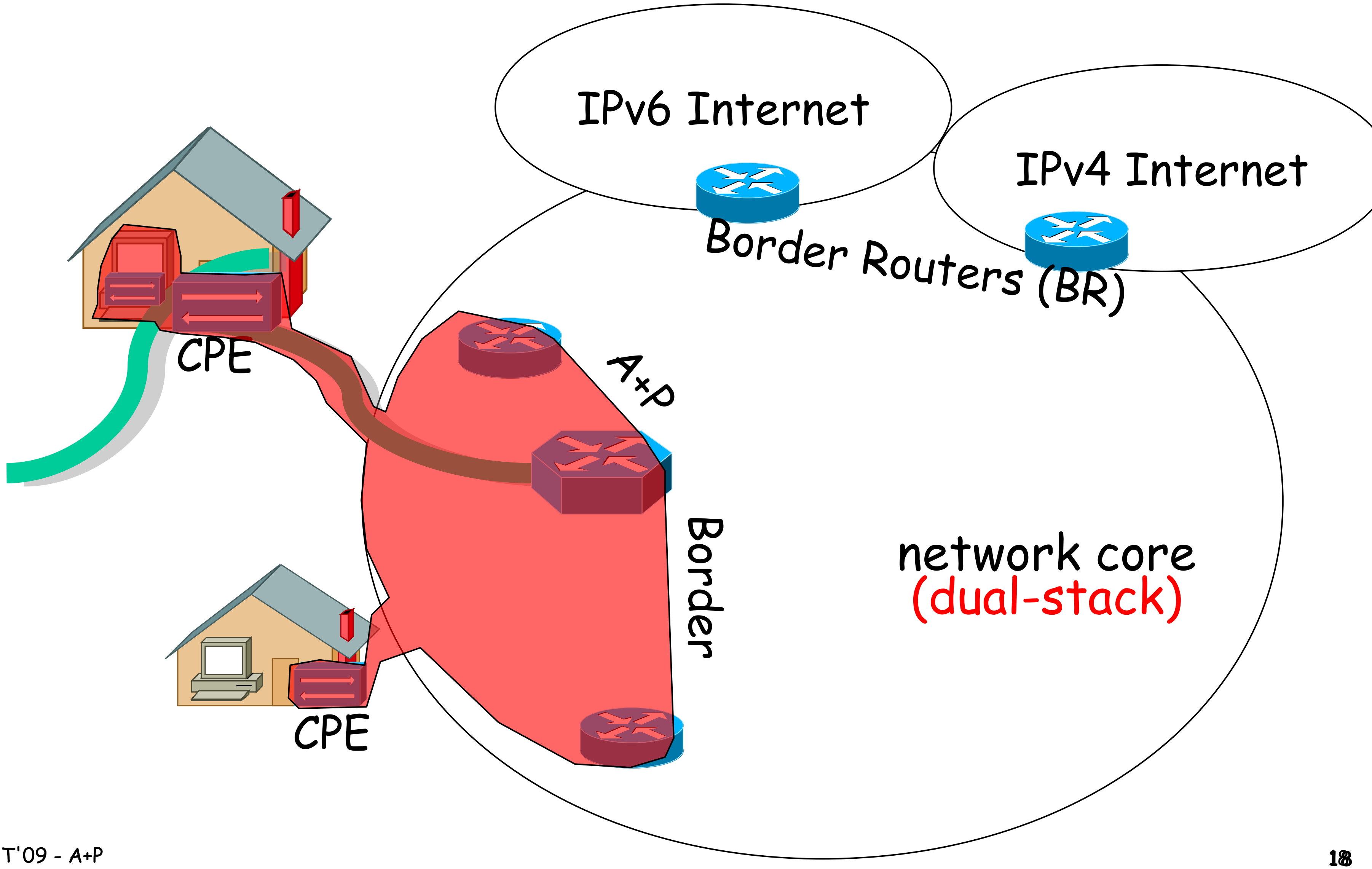
Constraints (II)

1) **"Double-NAT" shall be avoided.** Based on Constraint 2 multiple gateway devices might be present in a path, and once one has done some translation, those packets should not be re-translated.

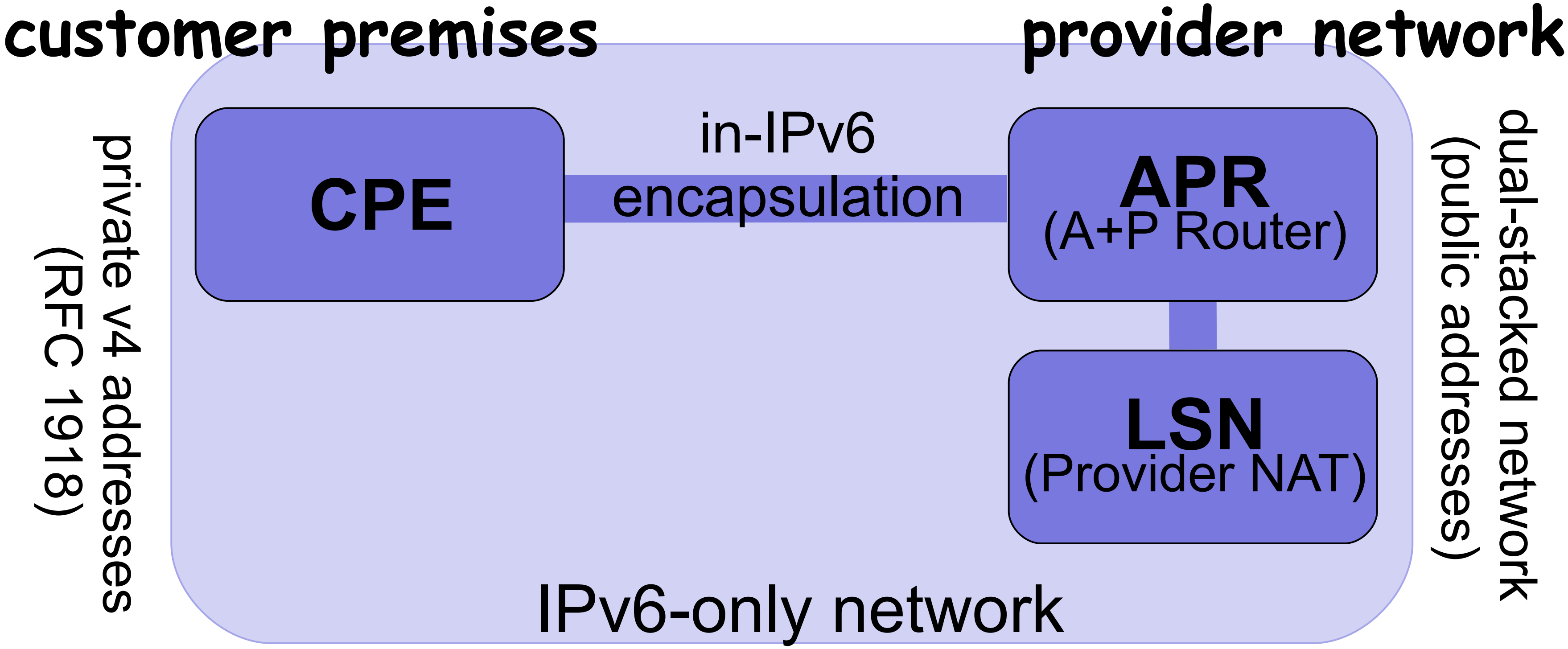
2) **Legal traceability.** ISPs must be able to provide the identity of a customer from the knowledge of the IPv4 public address and the port. This should have the lowest impact possible on the storage and the ISP. We assume that NATs on customer premises do not pose much of a problem, while provider NATs need to keep additional logs.

3) **IPv6 deployment should be encouraged.**

A+P Subsystem

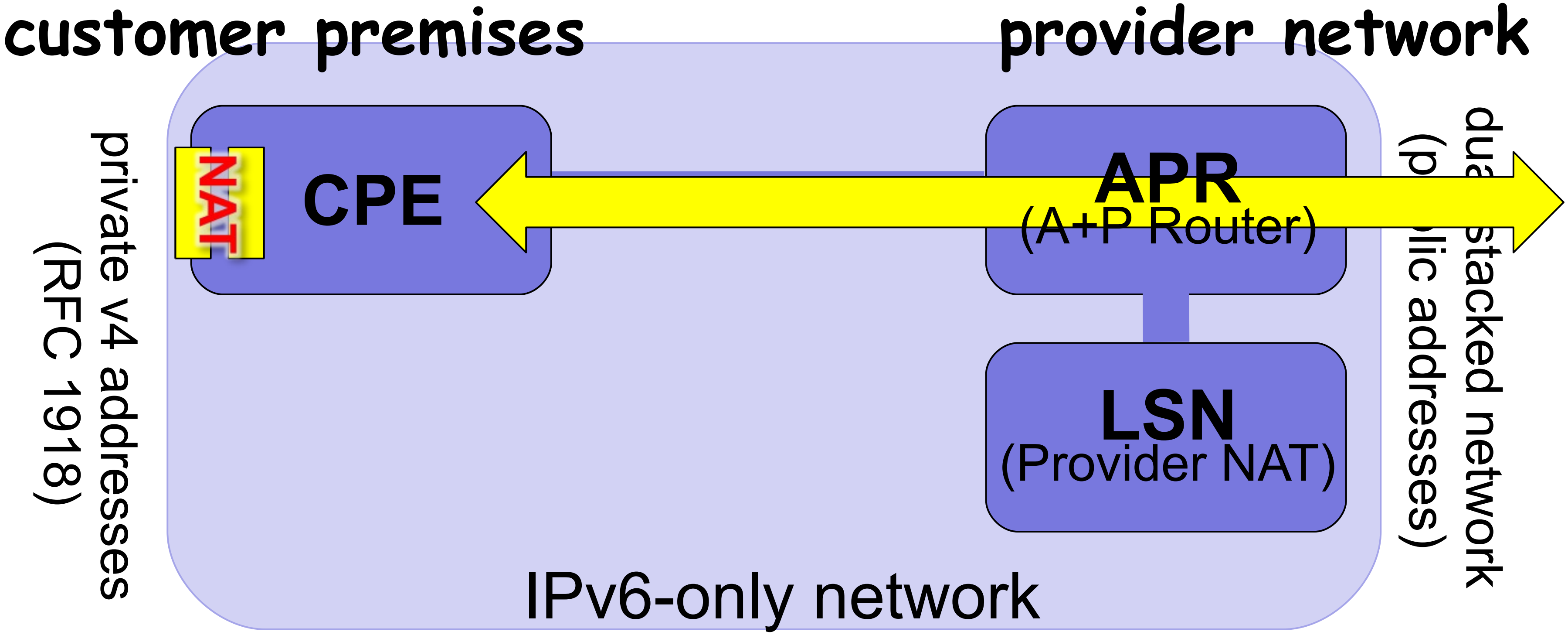


A+P Subsystem



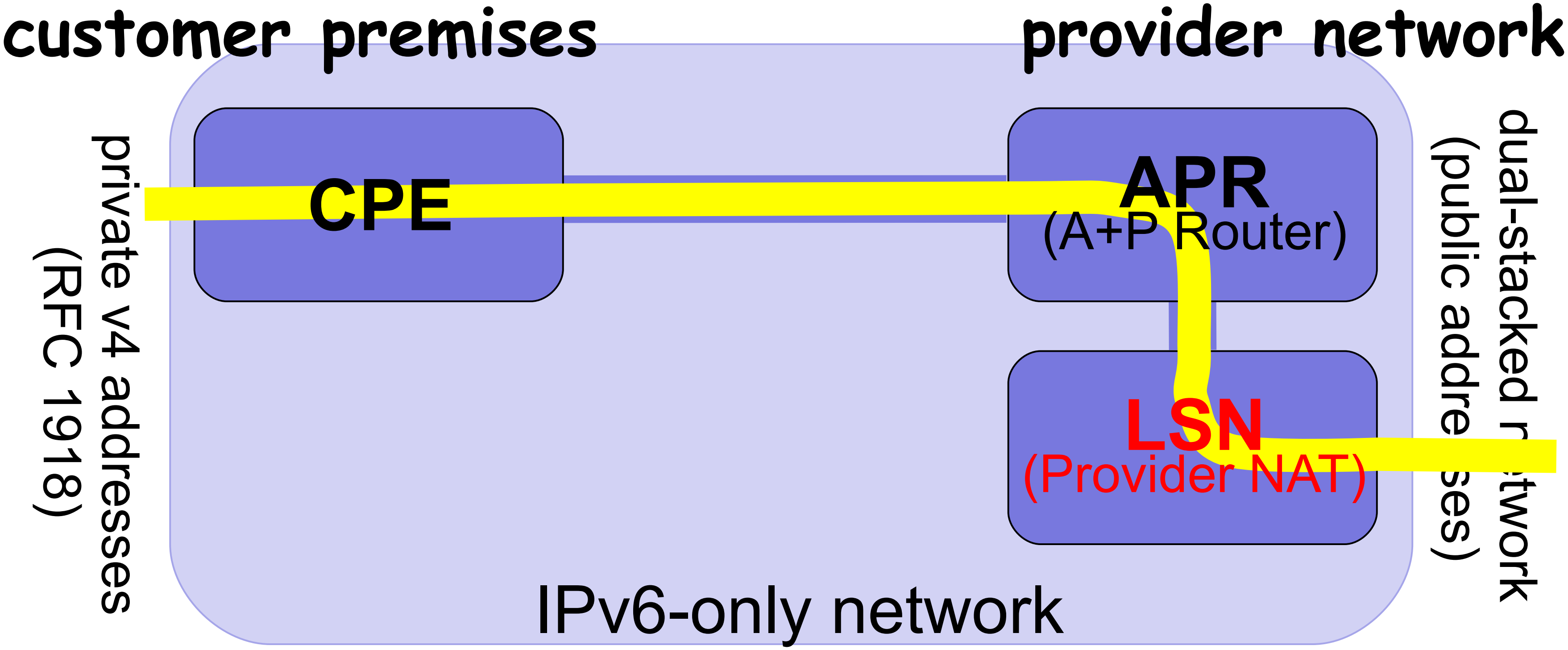
- "A+P pkts" are encapsulated in IPv6

A+P-NAT at CPE



- Untranslated end-to-end to CPE
- CPE nats to connect legacy hosts.
- APR encap/decaps only (LSN bypassed) !

Out-of-port-range pkts



- NAT could also be done at LSN
- However, customer has choice where NATing shall be done!

Port Allocation/Mapping

- Every customer may provision a fixed set of A+P ports, which are not touched. Users have services!
- Can manually specify reserved and mapped ports, e.g. via a web site that might look like a home NAT today
- A larger pool will be allocated on-the-fly that passes through LSN
- CPE learns IPv4 port restricted range via DHCP or other signaling cooperatively withxs LSN
- If LSN gets outbound NATted packet, it passes it after BCP38 check
- If packet is not NATted, LSN NATs it

Efficiency of Port Allocation

- Depends on traffic mix, but same as LSN
- A few inbound ports are typically used
- Outbound address compression ratio can be much higher due to heavy-tailed nature of traffic mix
- Actual measurement studies show typical port use is ~100 and peaks at 700

Status

Large router vendors are currently prototyping this functionality so that we can learn more through actual deployment exercises vs specification by committee

Open Questions

- Signaling mechanisms
- Port restrictions
- Assigned ports and IPv4 address
- Tunnel address of LSN

and your questions...