A+P Address Hack

The Revenge of the Stupid Core

Montréal / 2008.10.01

Randy Bush <randy@psg.com>

Olaf Maennel <olaf@maennel.net>

Luca Cittadini <luca.cittadini@gmail.com>

Steve Bellovin <smb@cs.columbia.edu>

<http://archive.psg.com/081001.A+P.pdf>

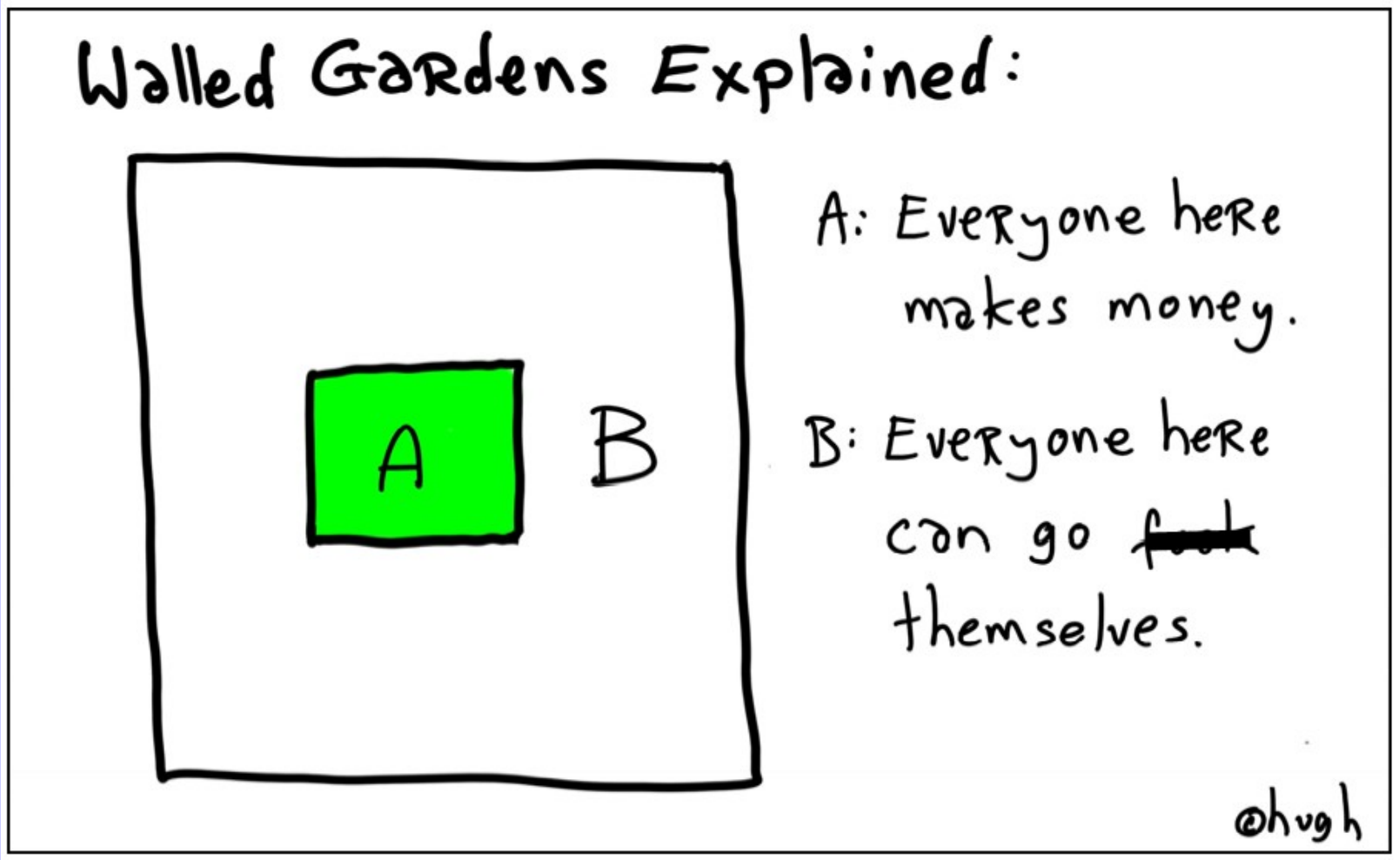
Problem Statement

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

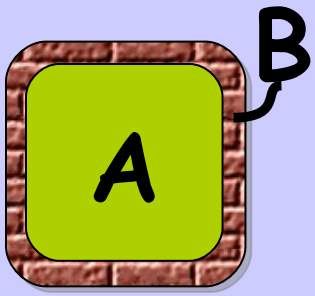
CGN Breaks the Net

- Not only does this cause problems for the carrier, but also for the whole net, as these captive customers can not try or use new disruptive technology
- NAT in middle of net has the problems of a smart core
- Walled gardens here we go!

I Googled "Walled Garden"



Walled Garden Re-Explained



C = The Global Internet
E.g. My Customers

A: Isolated, exploited, & restricted

B: Everyone here makes money

C: Everyone here can go fsck themselves

This Need Not Be Inevitable

Move the NAT to the CPE

As Alain 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."

If You Can't Roll CPE

- If you can not roll CPE immediately
- Then run a dual stack core
- The legacy CPE has a legacy IPv4 address now, let it keep it
- No need to break the Internet

A+P in One Slide

- Do the work at the CPE so that the **customer may control their fate**
- 'Steal' bits from the port number to extend the IPv4 Address
- Encapsulate in IPv6 in the ISP core and use normal routing to the edge
- Border Routers also en/decapsulate

"But This is Like X"

Nothing New Under Sun

- Late ARPANET ran out of address space with NCP circa 1981
- Needed to add more institutions
- Thus a *long leader* address extension
- No one wanted to rewrite kernels
- Greg Noel 'stole' unused short leader numbers and translated

A+P CPE is Modified

- Configured to use a restricted range of ports
- Configuration can be as simple or complex as you want it to be :)
- Some port bits dedicated to address extension, A+P
- NATs internal IPv4 to external A+P and encapsulates in IPv6

IPv6 Encap from CPE

- WKP = well known prefix, 4666::0/64
- Source of v6 packet is WKP+A+P
- Dest address of v6 packet
  - WKP+v4dest
- Border (BR) makes global v4 packet
  - source = A+P
  - dest = v4dest

Note That

- Normal IPv6 backbone routing is used
- Routing out from CPE is based on real destination, not pre-configured tunnel
- Only CPE and Border Routers are hacked
- No new equipment is introduced
- BRs do not have state or scaling issues

IPv6 Encap Toward CPE

- BR receives IPv4 packet w/ src/dest
- Encapsulates in IPv6 packet
  - src = WKP+src
  - dest = WKP+dest
- But note that dest is A+P
- It routes normally within ISP core

What Changes

- CPE - NATs and handles IPv4 A+P de/encapsulation in IPv6
- Border Router - de/encapsulates
- If you want to get into the kink of variable and/or dynamic length(P) games, life gets complex
- No extra hardware required

Transporting IPv6

- If the backbone is IPv6 capable, then IPv6 packets just move end to end
- If the backbone is not IPv6 capable, then the host or the site CPE must encapsulate to a 6to4 gateway or some other kink
- Deploy IPv6, it's forward not sideways

In an IPv4-only Core

- CPE sends packet with
  - Source of A+P
  - Dest of global IPv4 destination
- Outbound routes perfectly normally
- Replies need to be tunneled as they need to route A+P for the last mile
- Let's not go here

Nomenclature

It might be helpful to differentiate

- *Tunnel* goes from A, through some cloud, to B, i.e. has a predetermined end point, often pre-configured
- *Encapsulation* has no fixed end point, but goes from A, through the cloud, using normal IPv4/IPv6 routing, to an end point which is not predetermined

Thanks To

Dave Ward, for review, endless criticism, and questions