Issues with Port-Restricted IPs

Dave Thaler
dthaler@microsoft.com
Scope

• There are multiple types of issues:
  a. Inherent in address sharing (e.g. same as NAT)
     • covered in draft-ford-shared-addressing-issues
  b. As (a) but made worse with port-restricted IPs
  c. Specific to port-restricted IP addresses

This is scoped only to types (b) and (c)
General Issue

• Definition of “unicast address”
  – An identifier for a single interface
    • (within the scope: global, RFC1918, or link-local)
• Port-restricted IP’s change the definition so that multiple interfaces within the address’s scope get assigned the same address
• This is a change to the IP model as big as, but quite different from, the introduction of NAT
Implementation: hosts (1/2)

• Legacy IPv4 host IP stacks have no notion of port space limitation.
  – A+P cannot be deployed on legacy hosts
  – Will network refuse connectivity to customers?
• Even with an A+P aware kernel, applications expecting to bind to a specific port number will fail.
  – A difference from NAT is when the app sees a global IP it has no reason to believe anything is wrong
  – Proposed solution is to implement a NAT in the host kernel
    • Which means apps cannot communicate even with other on-link hosts without a NAT -> intra-link communication fails
    • And causes user confusion since default router is not the box they expect
Hosts (2/2)

• What about hosts with multiple interfaces?
  – E.g. app binds to IN_ADDR_ANY for on-link communication
    • Fails if you can’t get the same port on all interfaces
• What about hosts roaming between A+P networks and non A+P networks?
• How do the host IP stack and apps know how to switch back & forth between A+P and non A+P mode?
Management: ping

• ICMP msgs that don’t embed a packet (e.g. ping) have no ports
  – Customer initiated ICMP can be made to work with some effort, but not customer received ones

• In a pure A+P world, there is no way for a service provider technician to ping an A+P home router/host
  – If A+P is deployed on top of DS-lite, ping can be done over IPv6, but doesn’t provide liveness of IPv4 stack

• In contrast, ping etc. work fine within the area behind a common NAT
Other non-port based protocols

- The node assigned a port-restricted IP can no longer use non-port-based protocols even on the same link
- May not be a big deal when assigned to a home gateway if it doesn’t use any
  - But in some scenarios they might (e.g., pure A+P w/o IPv6 and gateway wants to do 6to4)
- But other hosts/applications/routers may
Provisioning system

• Service provider provisioning system would need to evolve to handle A+P
  – DHCP component
  – Databases
  – Management tools
  – Auditing/accounting, etc

• Those systems are complex and their evolution is costly and takes time

• This issue could be compounded by stateful dynamic port range allocation
Training/education

• Introducing such a complex change to the IP model requires retraining
  – Developers
  – Support personnel
  – Consultants
  – IT pros
  – Etc

• Again, this is over and above anything already inherent in address sharing
Security

• Port randomization is a security mitigation
  — draft-ietf-tsvwg-port-randomization
• Reducing the port space available to an application has negative security implications
• This issue is made worse if there is any port sub-delegation
  — Delegation hierarchy introduces wasted ports
Failure modes

• We already understand what fails with NATs and double NATs
  – many homes are already double-NATed today
• Port-restricted IP’s introduce lots of complexity with unknown (to most people anyway) failure modes
  – This will likely increase costs significantly compared to (say) multiple levels of NAT
Long-term impact

• Constant demand for IPv4 hacks to show up in IPv6
  – “We’re used to it”
  – “We want to do it the same way”
• Latest case in point: NAT66
• We don’t want this in IPv6
• We’ve learned that just saying “this is only for IPv4” doesn’t work
Summary

• Port-restricted IPs are a drastic change to the IP model
  – Lots of complexity
  – Lots of problems known, and probably more
  – People will get it wrong

• This architectural change is unnecessary
  – Multiple layers of NAT is already bad enough, this is arguably worse