Reducing Multicast in IPv6 Neighbor Discovery

draft-yourtchenko-colitti-nd-reduce-multicast

Andrew Yourtchenko

Lorenzo Colitti

Goals

- Evaluate, reduce(?) multicast traffic to hosts
- Config tweaks w/o changes to standards
- Introduce for discussion potential "minor" tweaks to standards

Problems that have been raised

- Multicasts on some links are slow / expensive
  - 802.11
- Multicasts impact battery life
  - No simple answer on whether better or worse than unicast
    - Different characteristics in different scenarios

How ND uses multicast

- DAD
  - 1 packet per IP address configured on the network
- RS
  - 1 packet per host that joins the network
- RA
  - Periodic: 1 packet every X seconds
  - Solicited: 1 packet for every host that joins the network
- NS
  - 1 packet for every new host/host pair

IMPACT

???

Does anyone have hard data?

It's not clear how much of a problem this actually is

MITIGATION

On-device multicast filtering

- Have firmware drop packets that don't go to groups the host has joined
- Problems:
  - Still requires receiving the packet
  - The depth of sleep may vary
  - Airtime inefficiency remains
- Implementation effort: host-local

Unicast solicited RAs

- **Respond to RS with unicast RA**
  - Standard already allows this

- **Substantial airtime and battery savings**

- **Should probably rate-limit and send multicast RA if too many RS/s**

```
16:09:20.901933 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:21.867637 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:23.599514 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:25.368509 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:27.039620 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:29.303073 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:30.899102 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:32.888595 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:35.186113 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:37.330651 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:39.781627 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:42.079485 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
16:09:43.333534 IP6 fe80::20d:48ff:fe0a:69cb > ff02::1: ICMP6, router
```

Infrastructure multicast filtering

- Perform multicast snooping
  - e.g., like SAVI does - RFC6620
- Convert multicast → unicast
  - Pure 802.11
  - Unicast ethernet (RFC6085)

Proxy ND on the Access Points

- Selectively translate multicast NS to unicast
  - Based on snooping L3/L2 mappings
- AP has all information it needs to do this
  - In 802.11 infra mode, traffic always goes through AP

Maximize interval for periodic RAs

- Set all lifetimes to current maximum limits:
  - AdvDefaultLifetime: 9000 seconds
  - MaxRtrAdvInterval: 1800 seconds

Increase Reachable Interval

- Makes hosts more tolerant to short failures
  - Avoid multicast NS when NUD fails
- Hosts can't learn about unreachable routers
  - OK if there is only one router or FHRP (VRRP, HSRP) pair

Clear on-link bit in prefixes

- Reduces NS multicasts from hosts
  - Everything non link-local is off-link

- Caveats:
  - By itself doesn't fix router->host multicast NS
  - Need to disable redirects on first-hop router
    - Conservative: prevent load on control plane, predictable debugging
  - All host-host traffic must go through router
    - If router is AP, or APs wired, not an issue

Explicit state with DHCPv6

- Store all state in the network:
  - Use DHCPv6 L3->L2 mappings
  - Drop ND between hosts
  - Router responds to ND on hosts' behalf
- Issues:
  - Some stacks do not support DHCPv6
  - Network in control of host address cycling
    - No host-generated privacy addresses, …

Client link shutdown when asleep

- Best savings (no power = no power drain!)

- Issues:
  - Additional delay on wake-up
  - No network-side wake-up
  - Breaks unicast, too

POTENTIAL STANDARDS TWEAKS

Raise max RA interval limit

- Remove AdvDefaultLifetime limit
  - Current limit 9000 seconds (Sender-only)
  - MaxRtrAdvInterval is another factor needing attention – current limit 1800 seconds

IETF89, London, Mar 2014

Explicitly client-driven RAs

- Explicitly client-driven RAs
  - When asleep:
    - Stop listening to multicast RAs
    - Retransmit RS when things expire
  - Retransmit RS when waking up
- Problems:
  - Breaks unsolicited RAs for asleep nodes
  - Slower / racy on wakeup
  - This is not a minor change