Scaling Layer 2

1000+ nodes

Linus Lüssing

June 2017

”What’s on second?”

© *Abbott and Costello: ”Who’s on first?”*

Outline

Outline

batman-adv

batman-adv

- A layer 2 mesh routing protocol

batman-adv

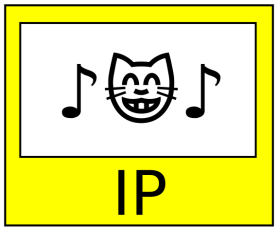
- A layer 2 mesh routing protocol
- Forwards frames

batman-adv

- A layer 2 mesh routing protocol
- Forwards frames
- Encapsulation

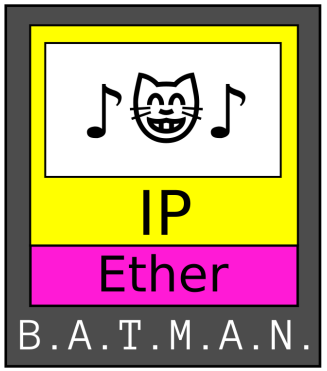
Encapsulation

https://www.flickr.com/photos/yourdon/3223467091
by Yourdon, CC-BY-SA



Encapsulation

https://de.wikipedia.org/wiki/Batmobil
Wikipedia, CC-BY-SA



Gains

Gains

- Network Protocol Agnostic

Gains

- Network Protocol Agnostic

Gains

- Network Protocol Agnostic
- Fast Client Roaming

Gains

- Network Protocol Agnostic
- Fast Client Roaming
- Network Coding, Bonding, ...

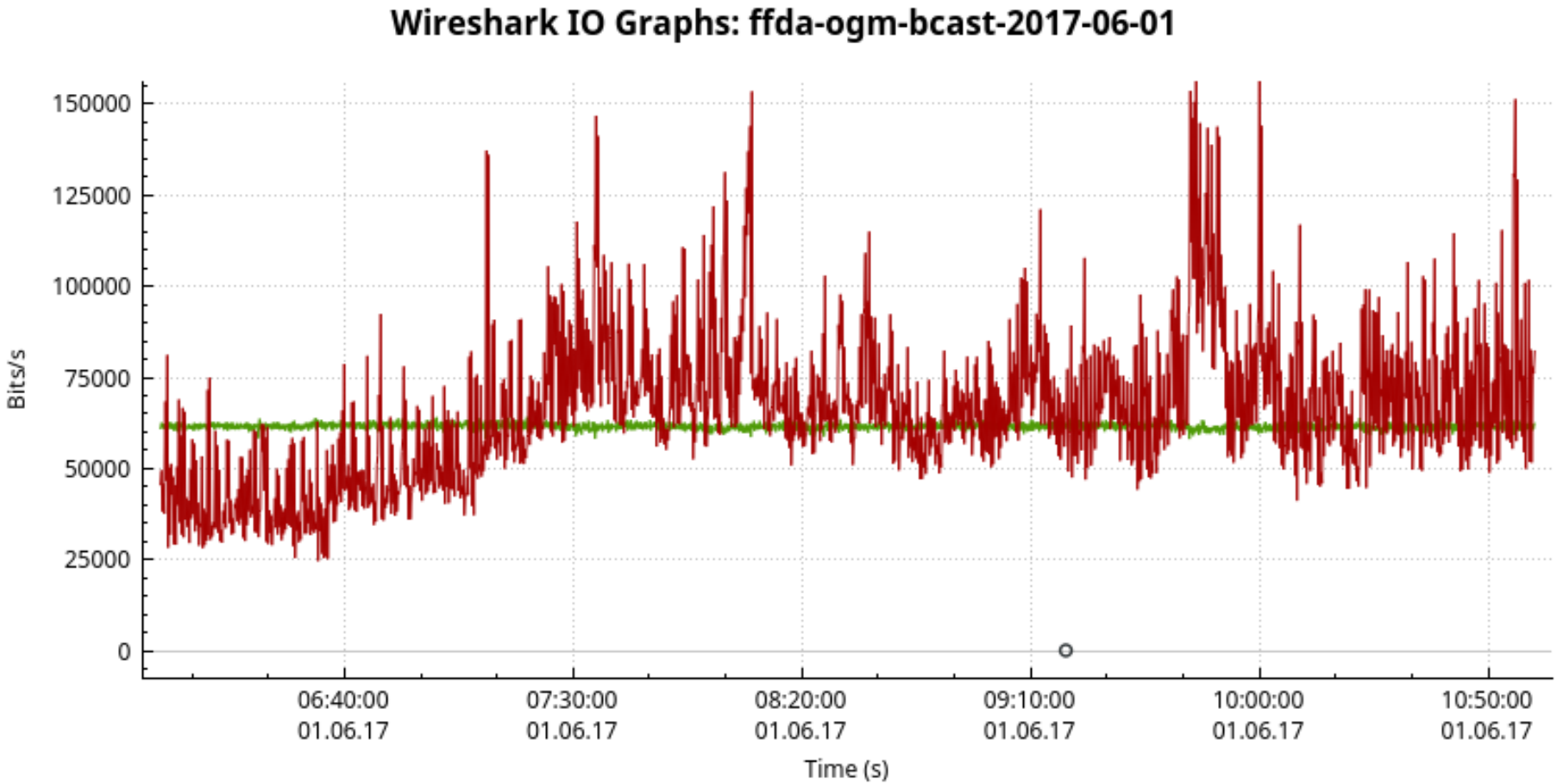
Gains

- Network Protocol Agnostic
- Fast Client Roaming
- Network Coding, Bonding, ...

Gains

- Network Protocol Agnostic
- Fast Client Roaming
- Network Coding, Bonding, ...
- Broadcasting

Costs



Freifunk Darmstadt:

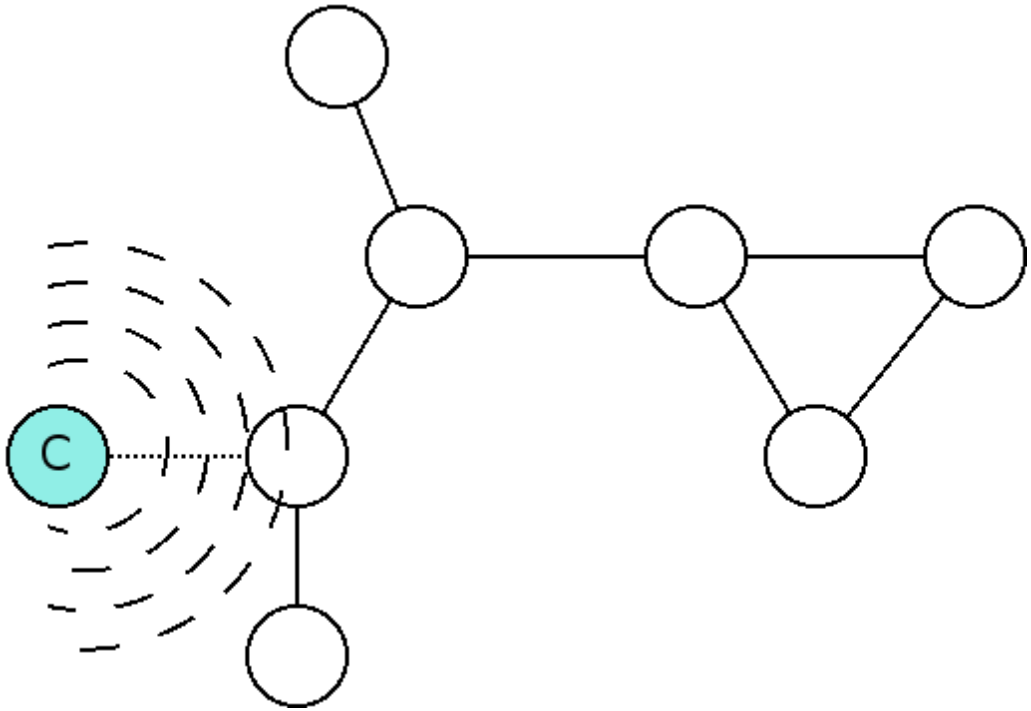
- ∼650 nodes
- ∼1500 clients

⇒

- OGM: 61 kbits/s
- Multicast: 66 kbits/s
- (Other: 0.066 kbits/s)

Gateway Feature

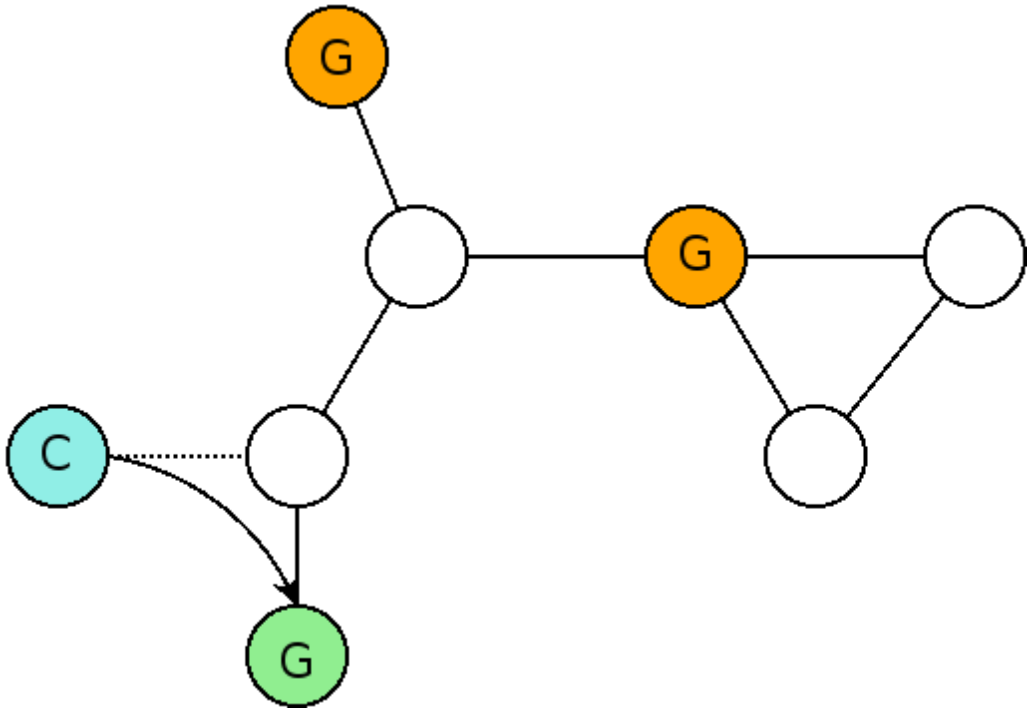
DHCP-Request/Discover



- Flooding

Gateway Feature

DHCP-Request/Discover



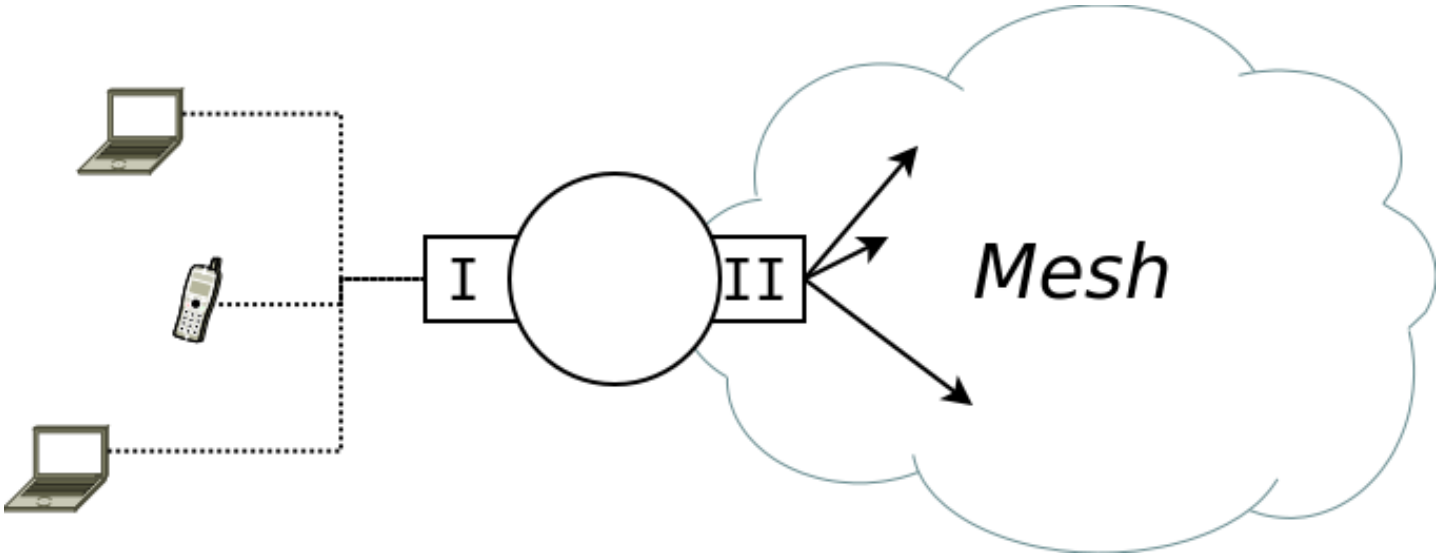
- Flooding ⇒ Unicasting

Distributed ARP Table

- Distributed Hash Table
- + ARP Cache

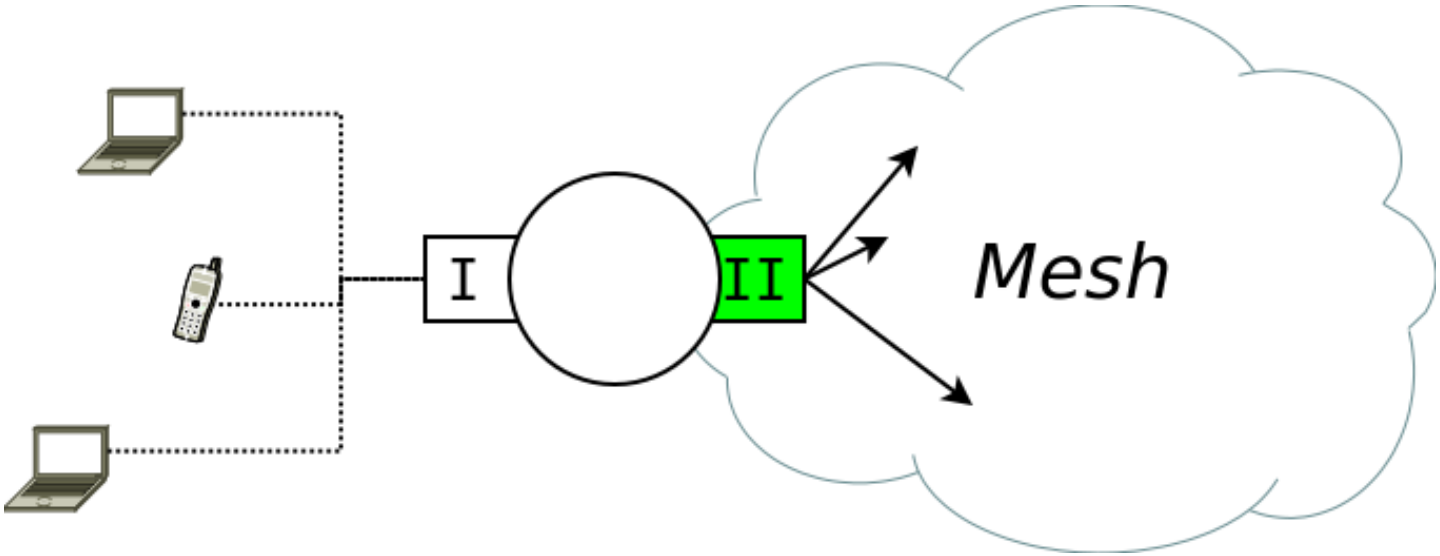
- Nice! IPv4 essentials tackled.
- What about IPv6?
- What about LLMNR/mDNS/...?

Gluon Architecture



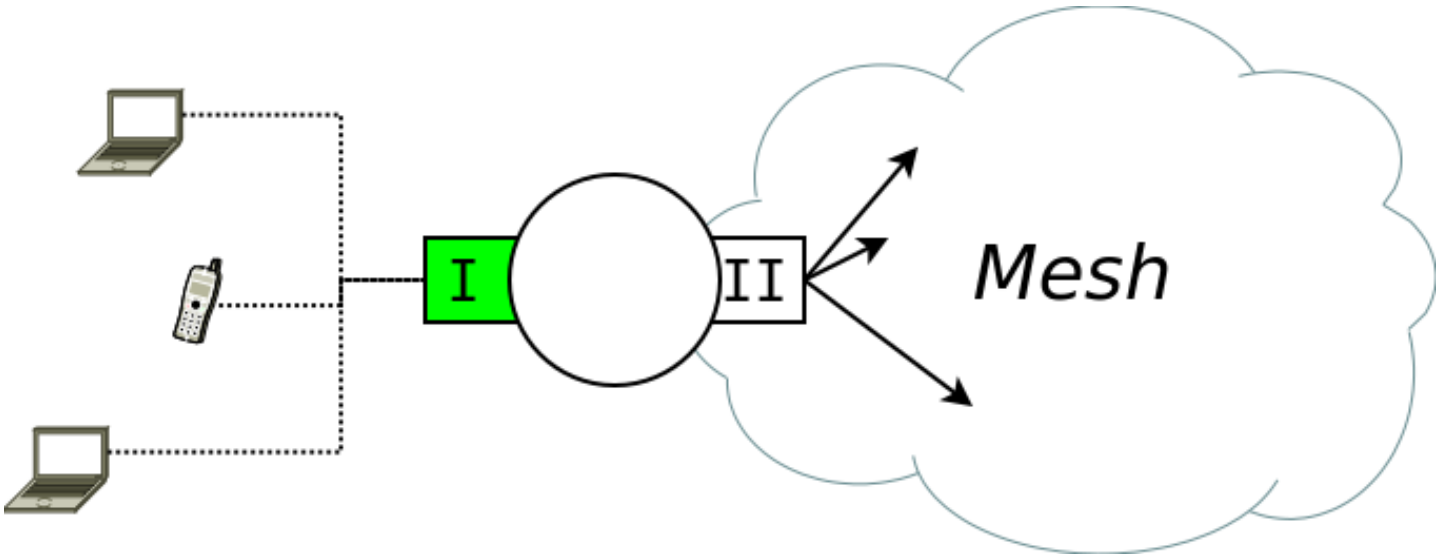
- (At least) 2 wireless interfaces:
  1. AP-Mode
     $\Rightarrow$ Client devices
  2. Adhoc-Mode or 802.11s-no-fwd
     $\Rightarrow$ Meshing

Rates & Intervals



@adhoc:

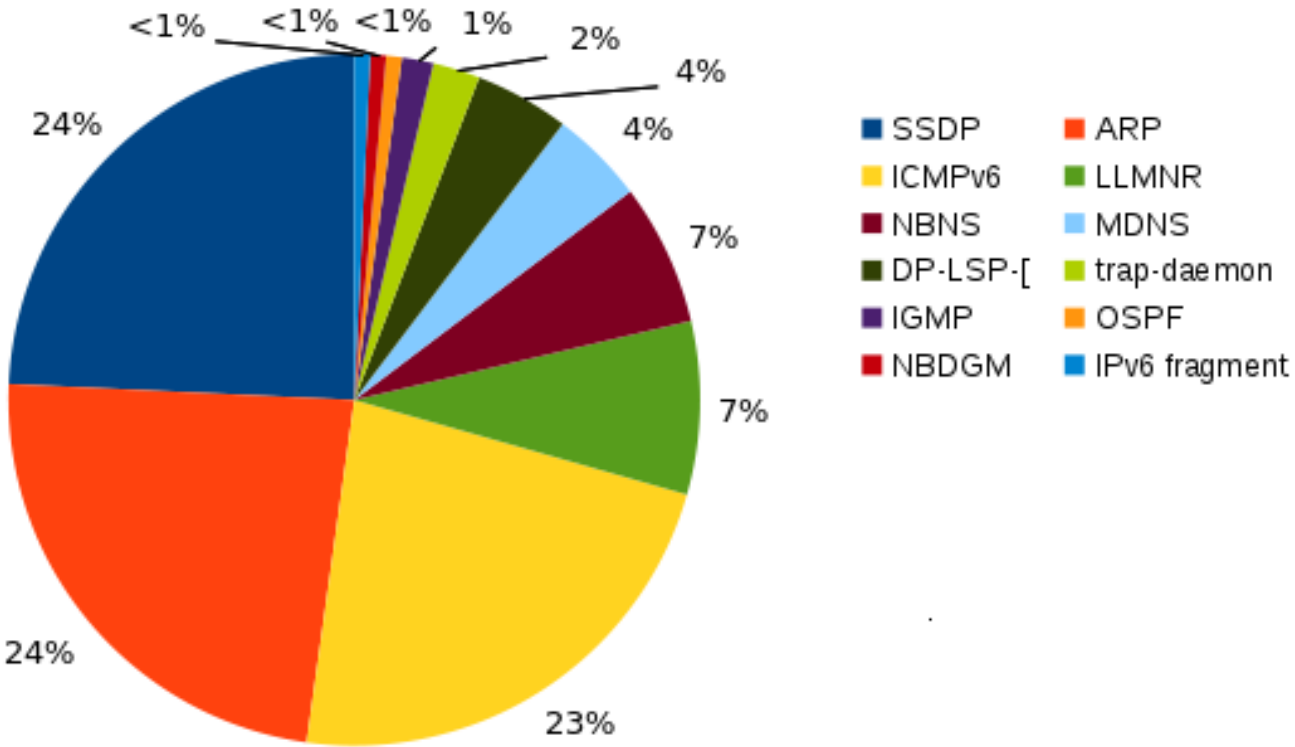
- Originator interval:
  1/s $\Rightarrow$ 5/s
- Multicast rate:
  1 $\Rightarrow$ 12 MBit/s (eff.: 0.5 $\Rightarrow$ 6 MBit/s)



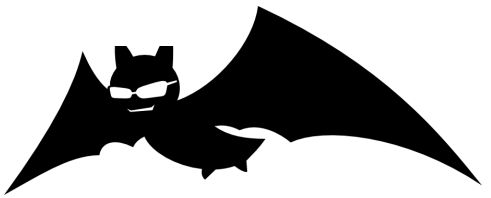
@ap:

- Removing supported rates:
  1 $\Rightarrow$ 6 MBit/s (eff.: 0.5 $\Rightarrow$ 3 MBit/s)

Layer 2 Overhead



Freifunk Lübeck, 2013



Multicast Filter

- Filter all multicast

Multicast Filter

- Filter all multicast

Multicast Filter

- Filter all multicast
- except:
  - ARP / DHCP
  - ICMPv6

ICMPv6

- Sometimes ”extra” ICMPv6 traffic

ICMPv6

- Sometimes ”extra” ICMPv6 traffic
- ”Ping of Death”: *ping6 ff02::1%wlan0*

ICMPv6

- Sometimes ”extra” ICMPv6 traffic
- ”Ping of Death”: *ping6 ff02::1%wlan0*
- ”Querier of Death”: MLD

ICMPv6: MLD

Multicast Listener Discovery (IPv4: IGMP)

- Used for:
  $\Rightarrow$ Multicast Routing

ICMPv6: MLD

Multicast Listener Discovery (IPv4: IGMP)

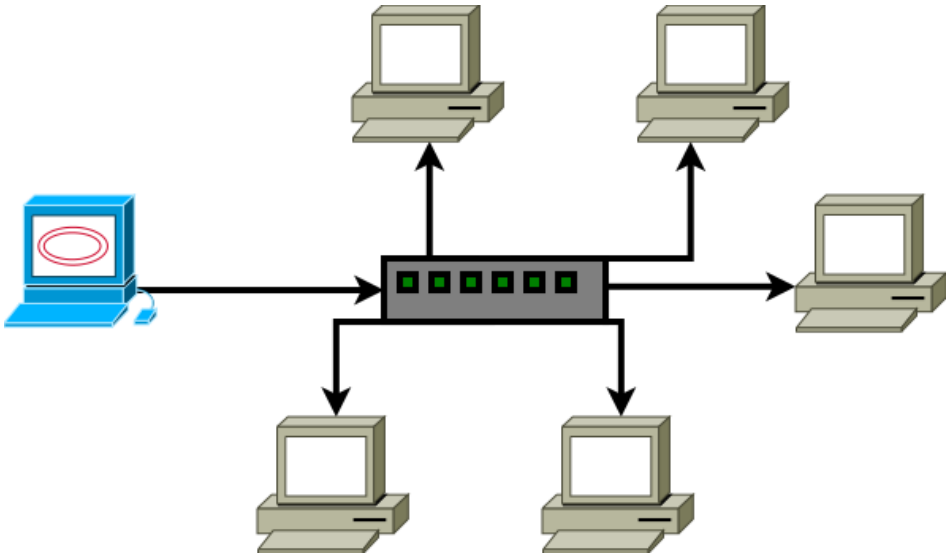
- Used for:
  $\Rightarrow$ Multicast Routing
- Multicast Routers query for listeners

ICMPv6: MLD

Multicast Listener Discovery (IPv4: IGMP)

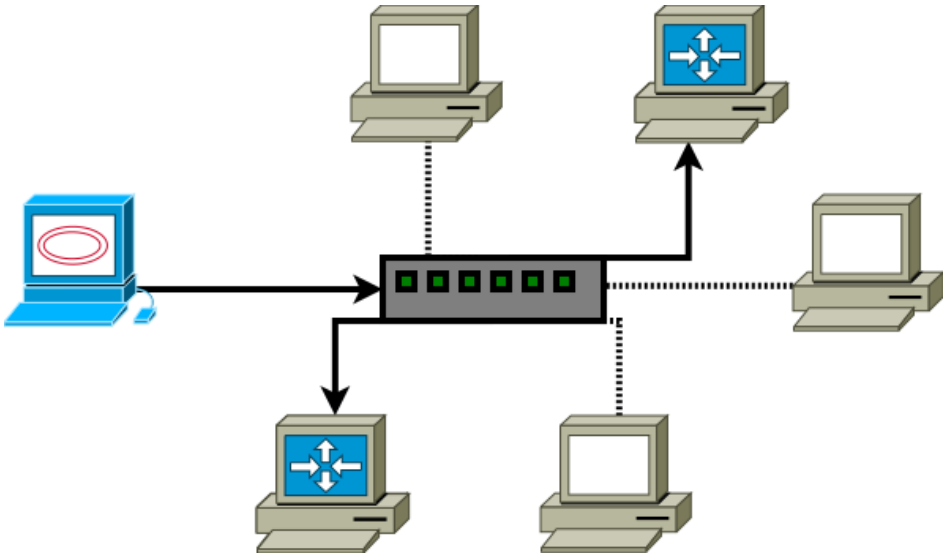
- Used for:
  $\Rightarrow$ Multicast Routing
- Multicast Routers query for listeners
- Multicast Listeners (receivers) respond

IGMP/MLD Snooping Switches



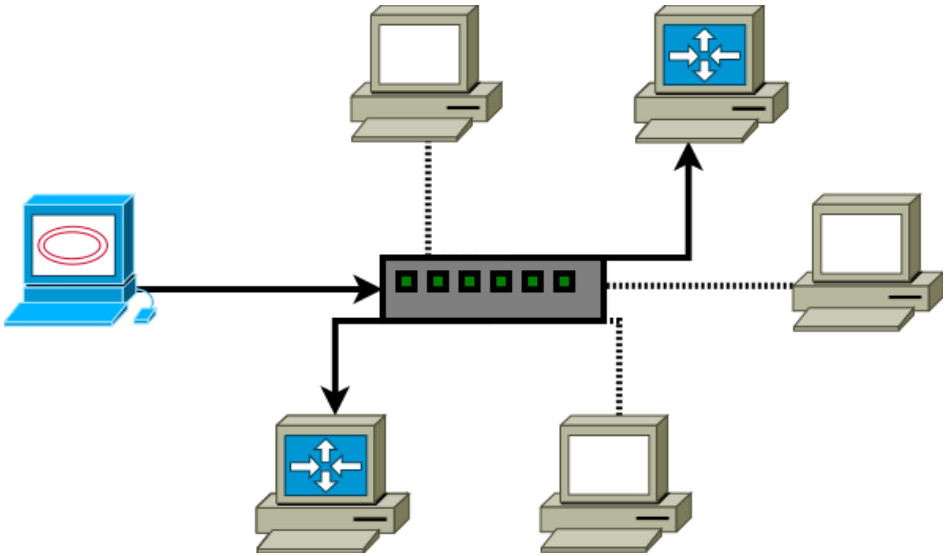
- Second use-case:
  $\Rightarrow$ ”Smart switches”

IGMP/MLD Snooping Switches



- Second use-case:
  $\Rightarrow$ ”Smart switches”

IGMP/MLD Snooping Switches

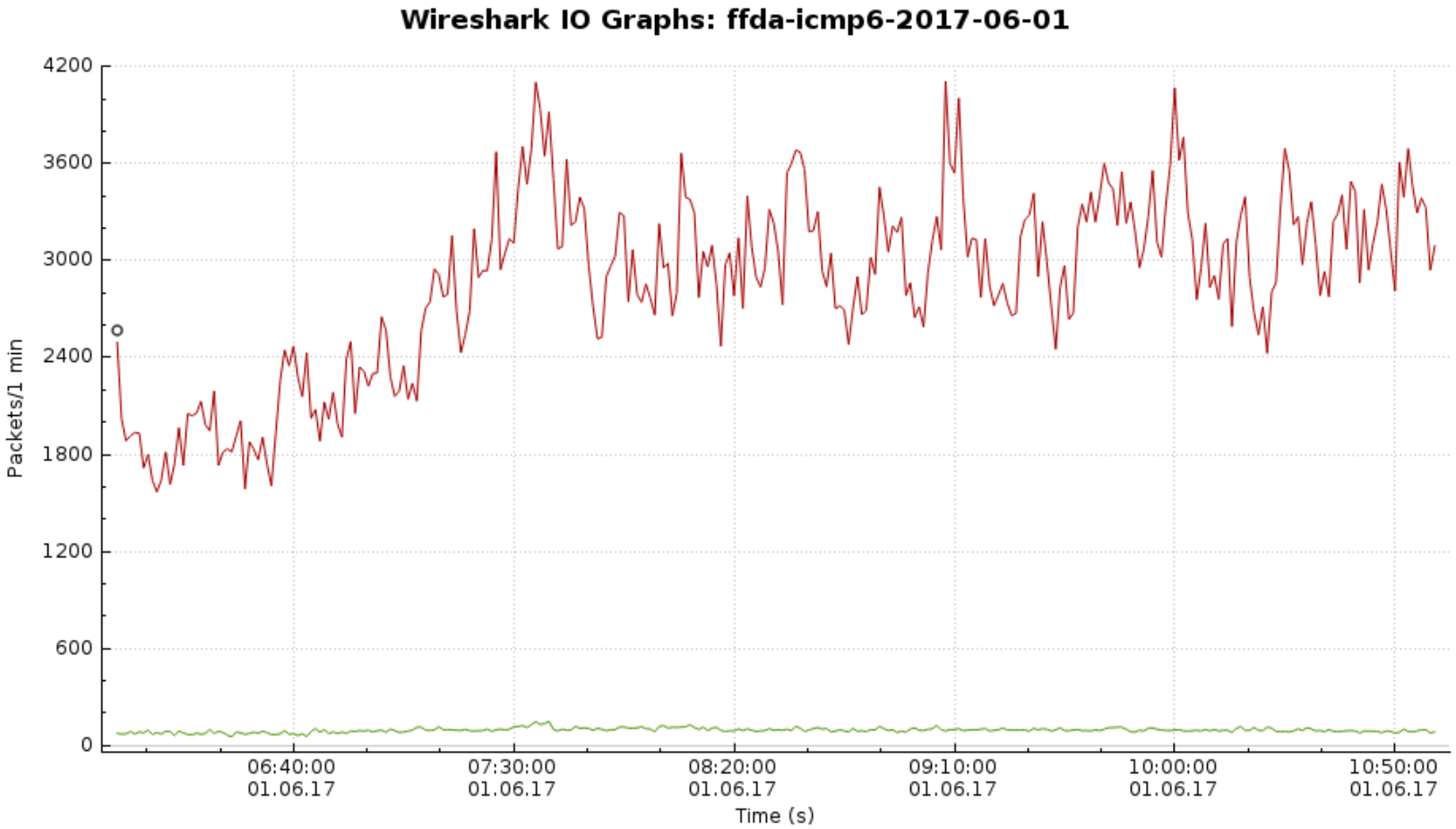


- Second use-case:
  ⇒ ”Smart switches”
- Software switch in Linux:
  ⇒ Linux bridge
  ⇒ Has IGMP/MLD Snooping support
  ⇒ Multicast-to-Unicast

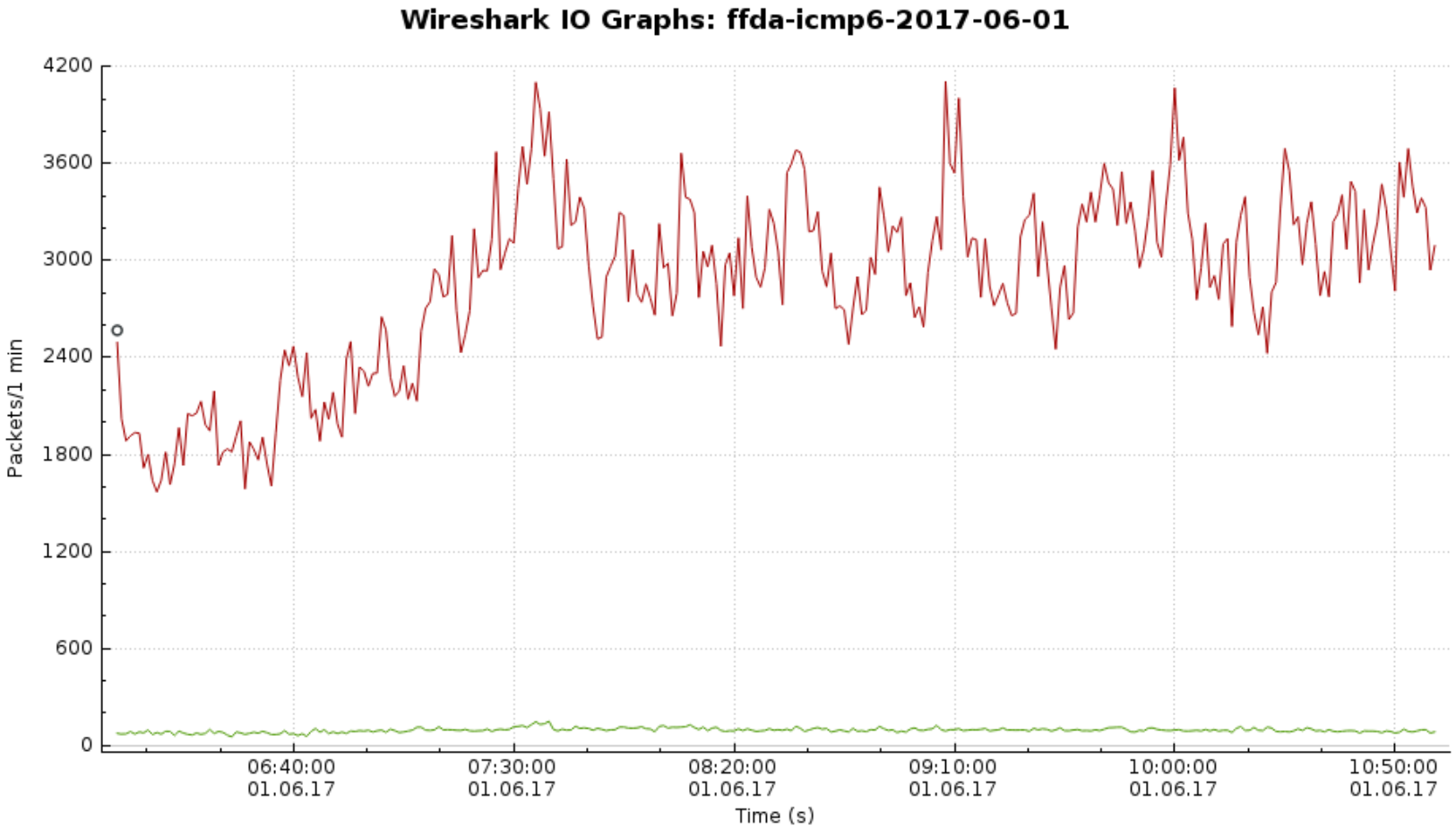
- Tricky to get right...

- Tricky to get right...
- Many bugs fixed in Linux

Impact for ICMPv6

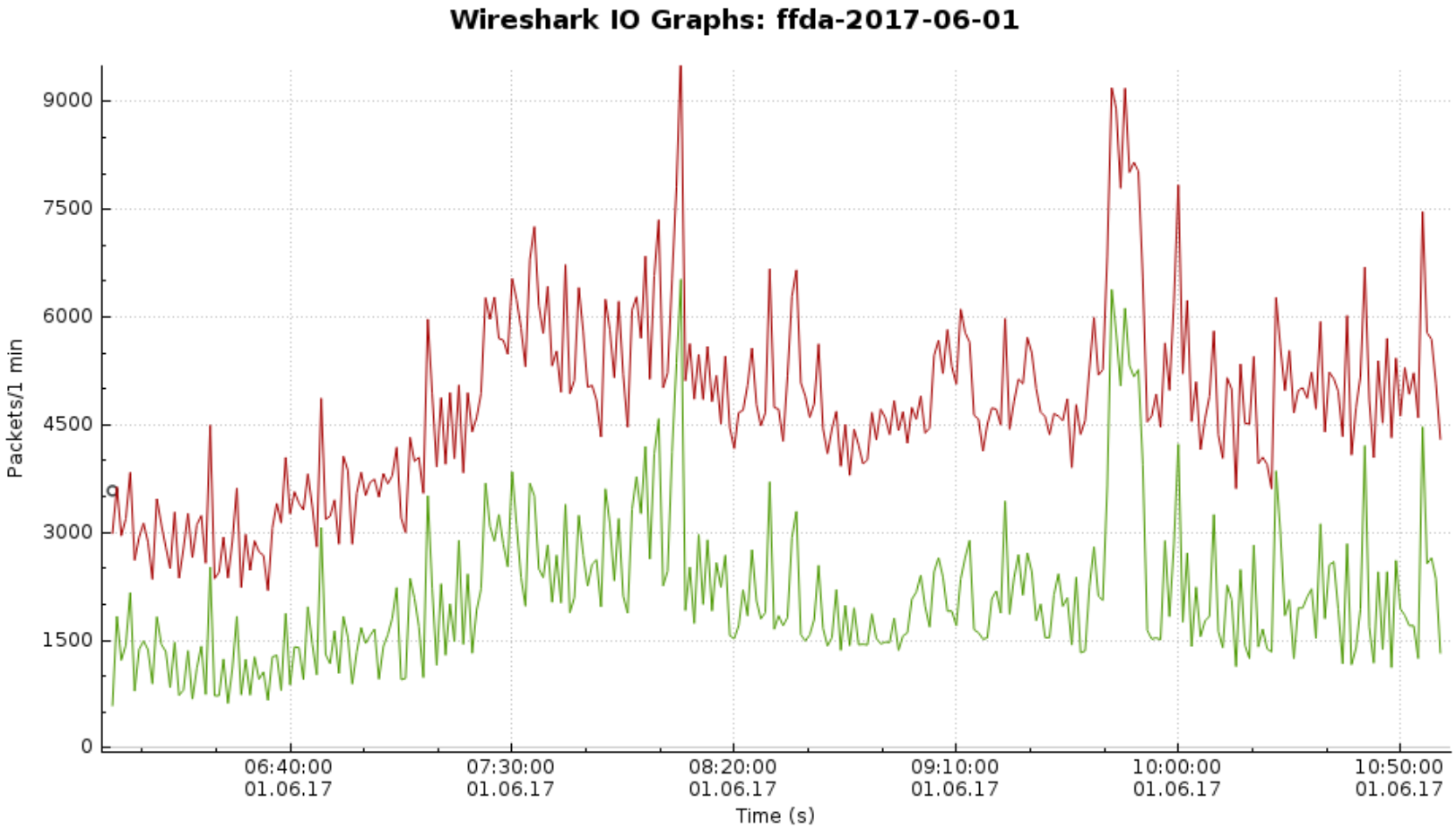


Impact for ICMPv6

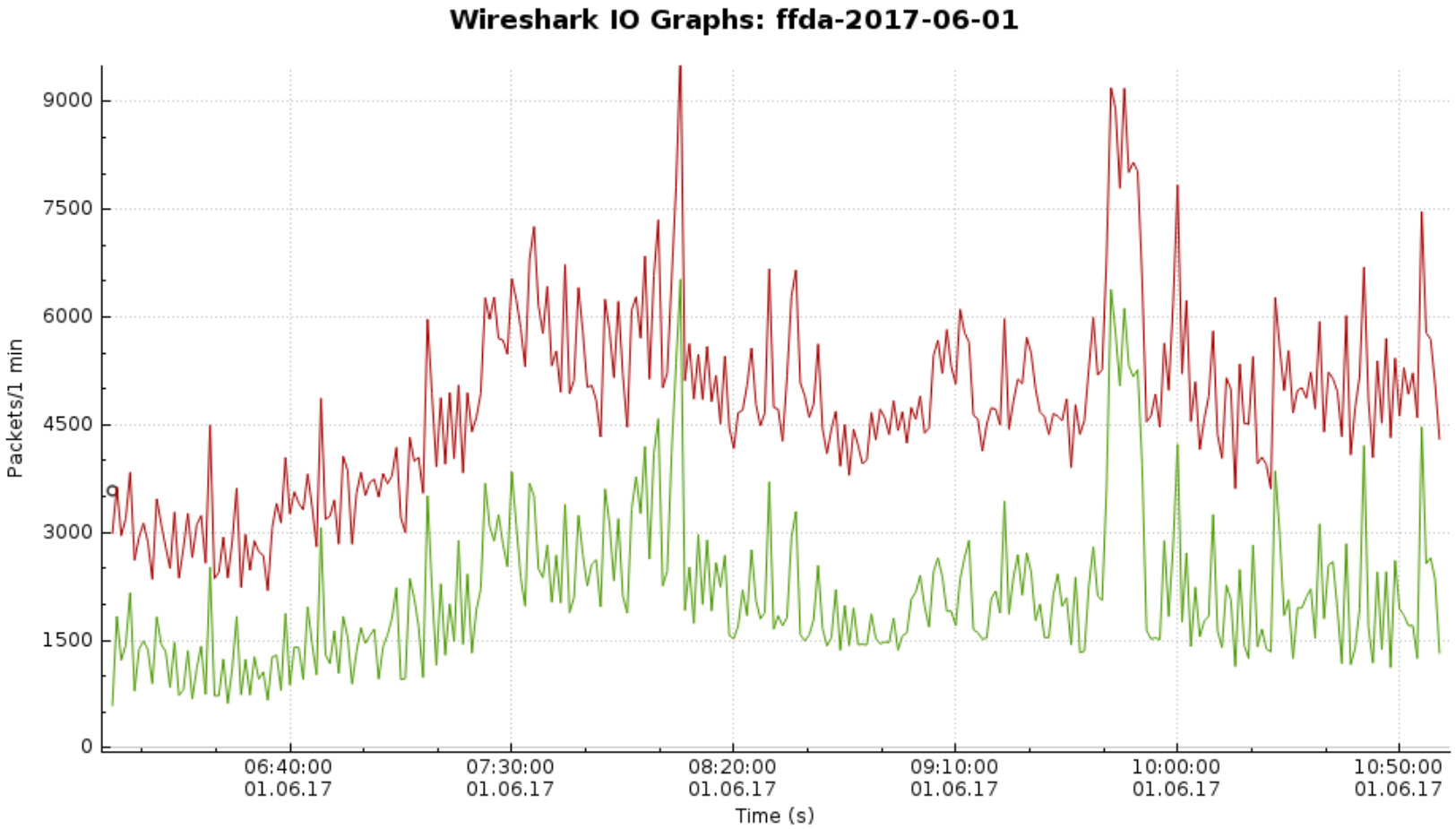


- *96.7%* reduction

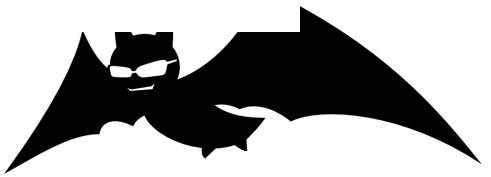
Impact total



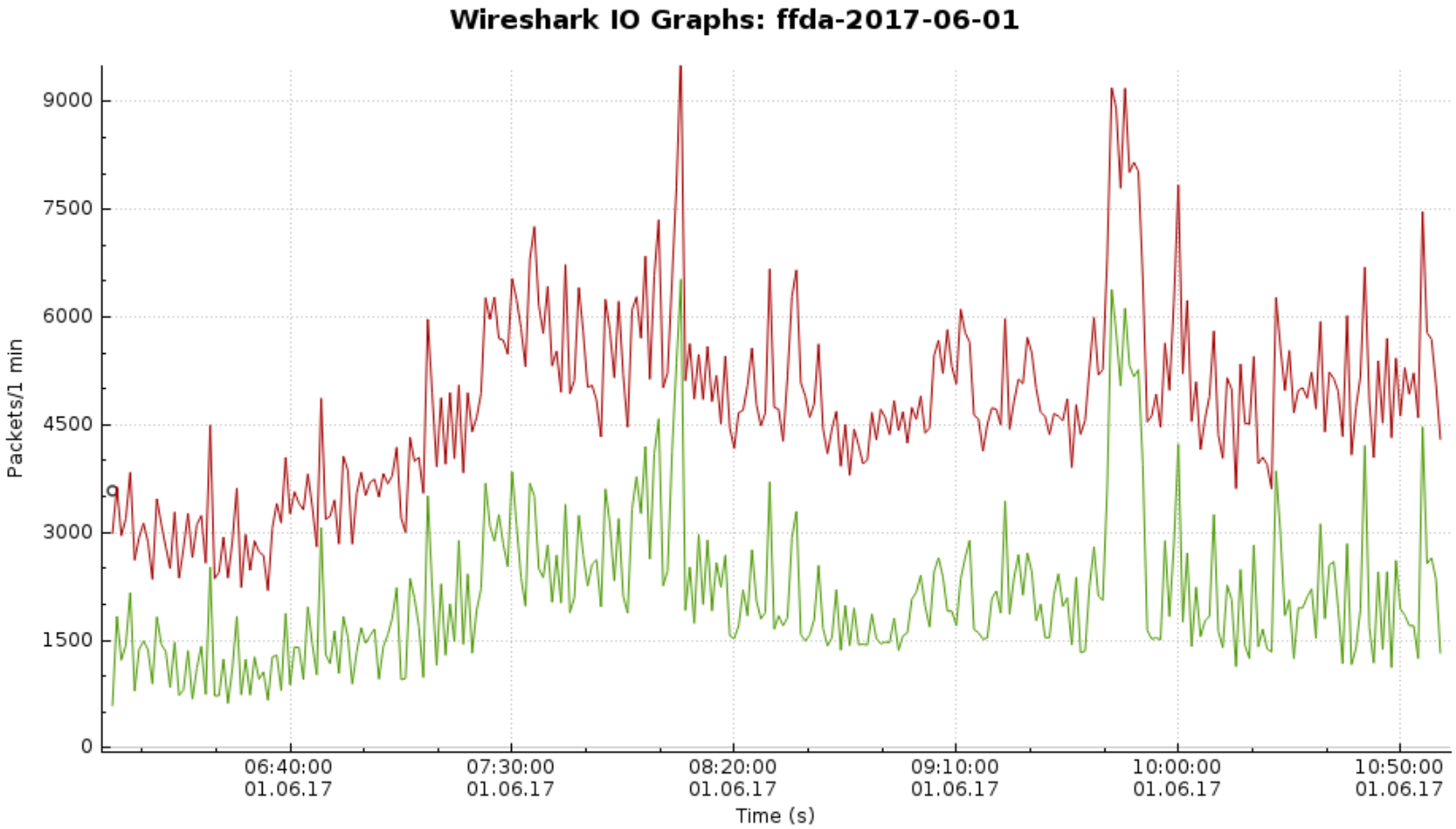
Impact total



- *55.79%* reduction



Impact total



- *55.79%* reduction
- Remaining: ARP (95.4%)

IPv6 Neighbor Discovery

Why? (IPv6 Neighbor Discovery vs. LLMNR/mDNS/...)

IPv6 Neighbor Discovery

Why? (IPv6 Neighbor Discovery vs. LLMNR/mDNS/...)

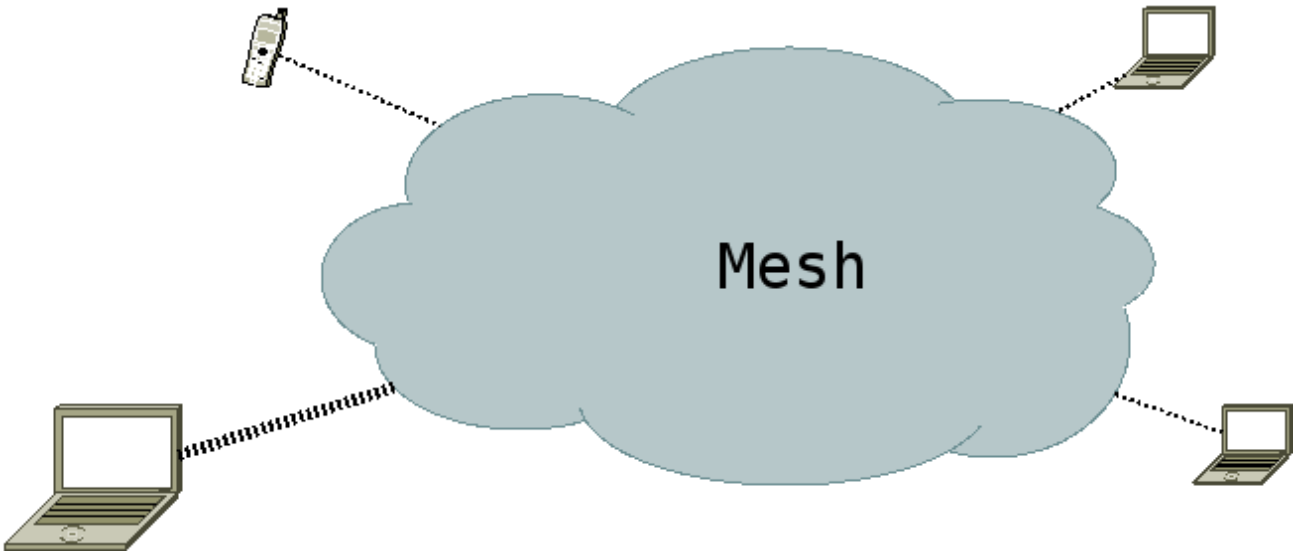
- Duplicate Address Detection:
  $\Rightarrow$ 0 multicast listeners

IPv6 Neighbor Discovery

Why? (IPv6 Neighbor Discovery vs. LLMNR/mDNS/...)

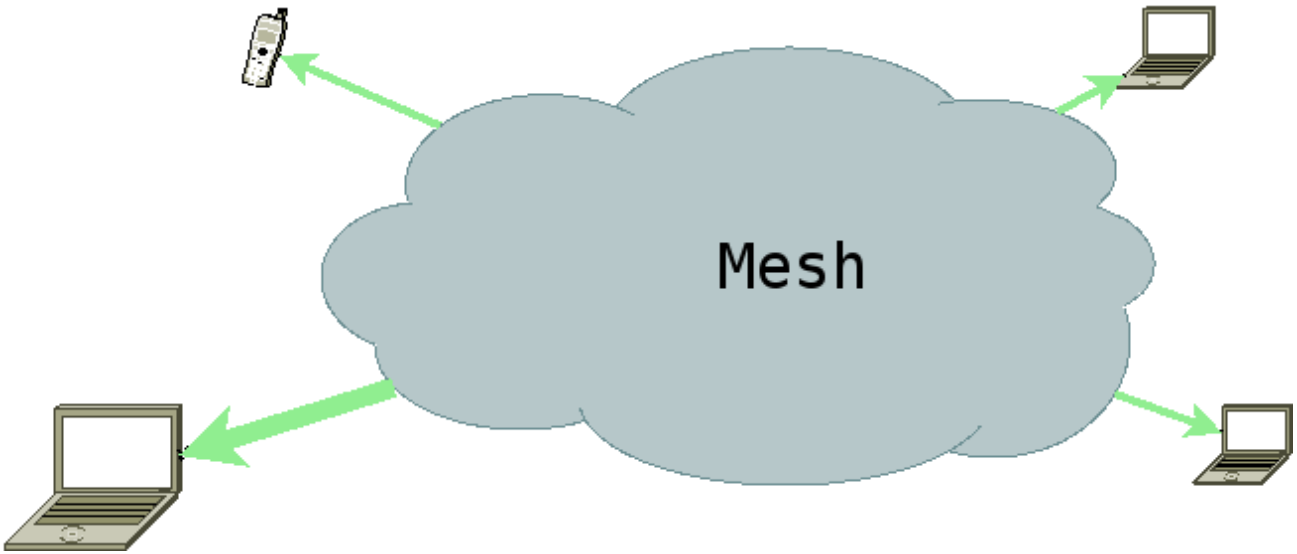
- Duplicate Address Detection:
  $\Rightarrow$ 0 multicast listeners
- Address resolution:
  $\Rightarrow$ 1 multicast listener

IPv6 Neighbor Discovery



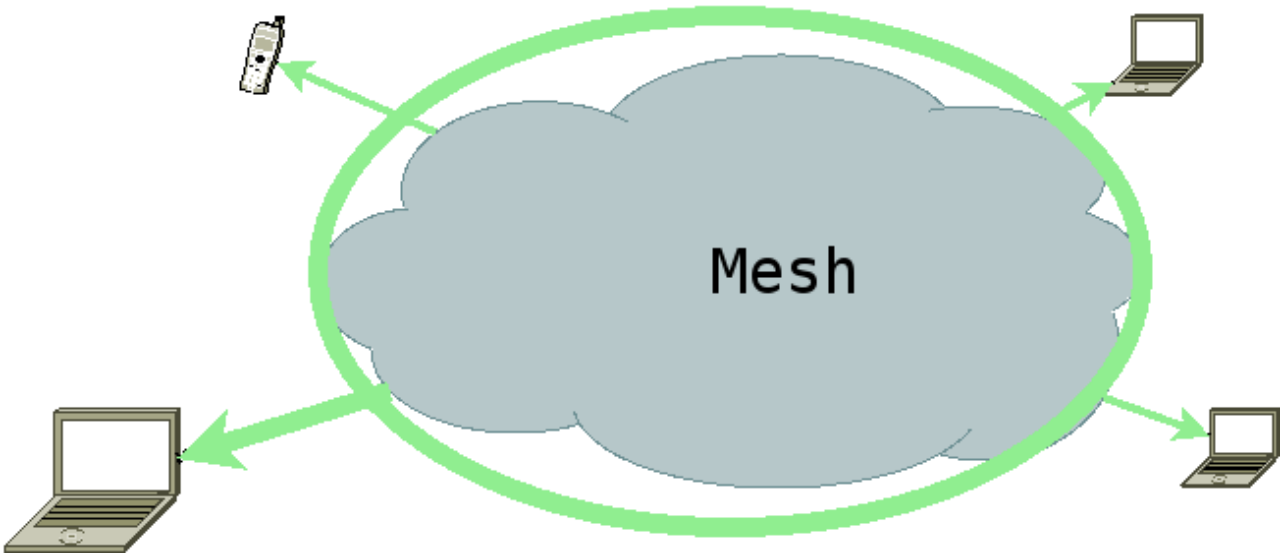
- Works well with snooping

IPv6 Neighbor Discovery



- Works well with snooping

IPv6 Neighbor Discovery



- Works well with snooping
- Extend to mesh?

Multicast-Aware Nodes

Since batman-adv $\geq$ v2016.3:

- batman-adv got multicast-aware

Multicast-Aware Nodes

Since batman-adv $\geq$ v2016.3:

- batman-adv got multicast-aware
- bridge learns from IGMP/MLD
  $\Rightarrow$ batman-adv learns from bridge

Multicast-Aware Nodes

Since batman-adv $\geq$ v2016.3:

- batman-adv got multicast-aware
- bridge learns from IGMP/MLD
  $\Rightarrow$ batman-adv learns from bridge

New forwarding rules:

- 0 listeners: $\Rightarrow$ Drop

Multicast-Aware Nodes

Since batman-adv $\geq$ v2016.3:

- batman-adv got multicast-aware
- bridge learns from IGMP/MLD
  $\Rightarrow$ batman-adv learns from bridge

New forwarding rules:

- 0 listeners: $\Rightarrow$ Drop
- 1 listener: $\Rightarrow$ via Unicast

- Graphs?

- Graphs?
- Not yet, part of upcoming Gluon release

Summary

- Layer 2:
  $\Rightarrow$ Noisy and wild (by default)...
- Layer 2 essentials:
  $\Rightarrow$ Optimized by batman-adv
- Layer 2 non-essentials:
  $\Rightarrow$ Filtered by Gluon

Summary / Next steps

- Further investigations regarding ARP
- Scaling next: Mesh Protocol (BATMAN V)
- Further utilizing multicast awareness (multicast streaming etc.)

- Questions?

- Thanks!