Mitigating Superfluous Multicast Data Traffic and Control State

John Kristoff
jtk@northwestern.edu
http://aharp.ittns.northwestern.edu
+1 847 467-5878
Northwestern University
Evanston, IL 60208

This is an IPv4 only talk

I know a lot less about IPv6 multicast, so I am conveniently ignoring IPv6 details and examples. I think I can safely say that some of the same issues apply, some do not and there may be some new ones.

Recent deployment statistics *

- ~450 ASNs representing ~7500 MBGP routes
- ~245M addresses covered by those routes
    - Real reachability is probably much lower
    - It says nothing about PIM-enabled interfaces
- ~1300 (not totally bogus) SAs being announced

* From http://www.multicasttech.com/status/

If you enable multicast, this talk matters because...

- Unless configured otherwise

  1. any host can send to any group
  2. any host can receive any group's traffic

- Flooding attacks should be obvious, but also...
- Hosts trigger additional network device state and processing workload
- Your unicast+mcast nets are coupled together?

Open receiver model areas of concern

- IGMP group state for each router interface
- *,G and S,G PIM state for each router interface
  - state created on upstream routers and RP
- L2 device must peek into L3 for IGMP clues
- L2 device maintains group-to-port table
- Does group's send rate exceed a receive path?

Open sender model areas of concern

- Router (DR) encapsulates and unicasts to RP
- Router maintains PIM state for each interface
    - state created on upstream routers and RP
- Packets replicated at branches
- MSDP SA state and SA flooding to peers
- MSDP receiver SA decapsulation effort
- SMURF-style amplification attacks

State, good news and bad news

- Good
  - Most state entries will time out after a short period of time (usually within just a few minutes) if the sender/receiver is inactive
- Bad
  - It doesn't take long to cycle through 224/4

What you probably don't need

```
(*, 239.255.255.253), 10w1d/00:03:25, RP
192.0.2.1, flags: SJC
  Incoming interface: Null, RPF nbr 0.0.0.0
  Outgoing interface list:
    Vlan1, Forward/Sparse, 1d13h/00:03:11
    Vlan2, Forward/Sparse, 5d12h/00:02:32
    Vlan3, Forward/Sparse, 2w3d/00:02:34
    Vlan4, Forward/Sparse, 2w3d/00:02:45
    Vlan5, Forward/Sparse, 2w3d/00:03:25
    Vlan6, Forward/Sparse, 2w3d/00:03:13
    ...
```

You don't need this either

```
border> show msdp sa | match 224.0.1.76
224.0.1.76    192.0.2.1  192.0.2.11 ...
224.0.1.76    192.0.2.2  192.0.2.12 ...
224.0.1.76    192.0.2.3  192.0.2.13 ...
224.0.1.76    192.0.2.4  192.0.2.14 ...
224.0.1.76    192.0.2.5  192.0.2.15 ...
224.0.1.76    192.0.2.6  192.0.2.16 ...
224.0.1.76    192.0.2.7  192.0.2.18 ...
224.0.1.76    192.0.2.8  192.0.2.19 ...
...
```

An overview of 224/4 usage

- 224.0.0.0/24 local control, link local only
  - Remaining 224/8 various assignments / uses
- 225/8 - 231/8 IANA reserved
- 232/8 source specific multicast (SSM)
- 233/8 GLOP space
- 234/8 - 238/8 IANA reserved
- 239/8 administratively scoped

Multicast scoping

- Scoping usually refers to data plane forwarding
- Limits the distribution of certain multicast traffic
- This is kind of like unicast bogon filtering
  - can't use a multicast bogon server today
- Scoping control plane packets is critical
  - we can scope that too, but it is a pain to do

Bogon filtering multicast space

- Instead of supporting up to /4 (~268M) groups
- we'd be left with slightly more than a /8 (~20M)
- or less than 10% of original Class D space.
- You can reduce this quite a bit more
  - but managing those filters can be a PITA!

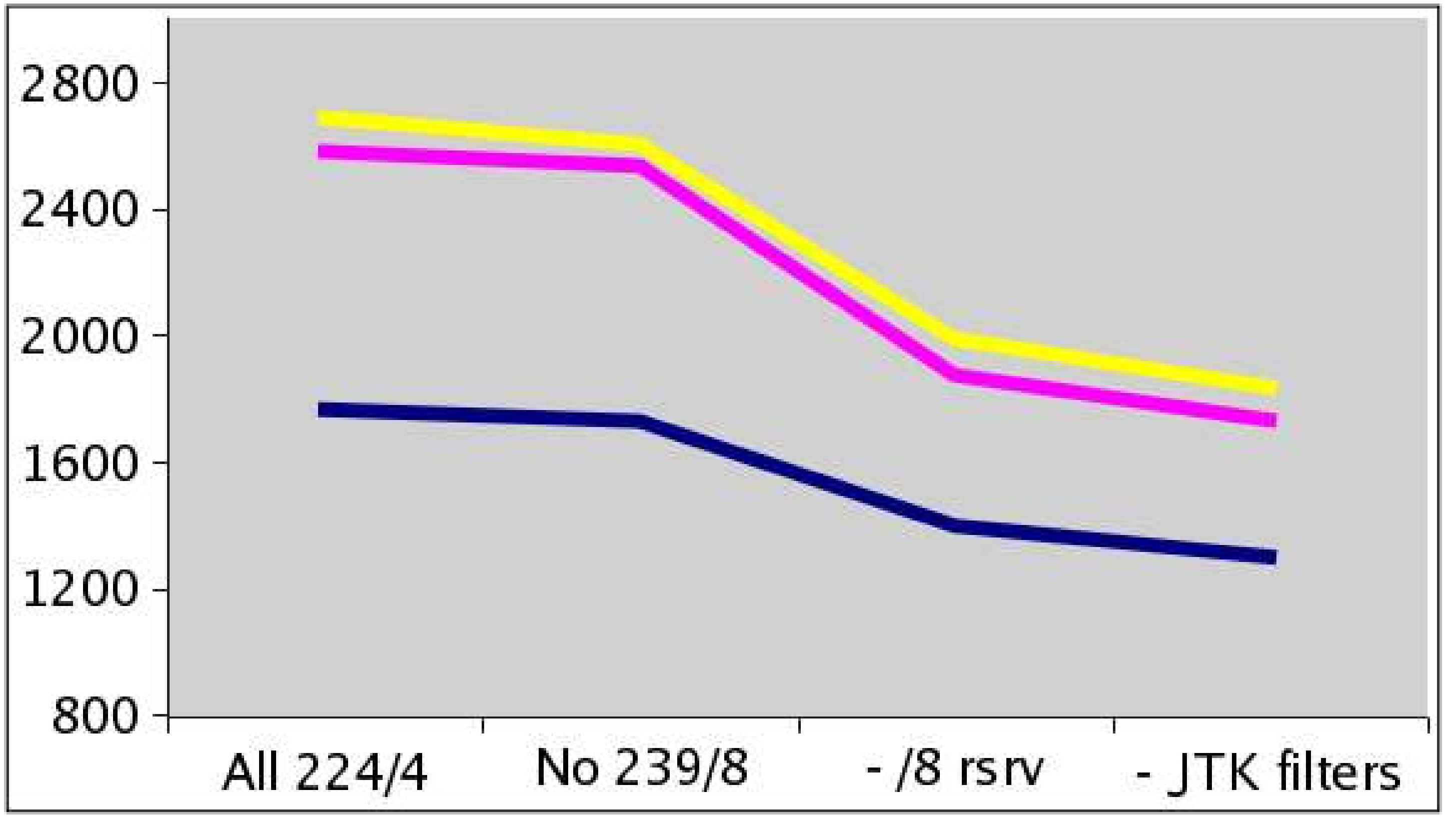
Why simple filtering doesn't help

- Still the potential for millions of MSDP SAs,
- and millions of PIM routes across routers,
- and millions of IGMP entries across interfaces,
- ...and BTW, multicast CIDR doesn't exist

Why simple filtering does help

It mitigates stupid unicast-based worms

Bogon filtering effects on today's global MSDP SAs



Juniper data plane scoping example

```
routing-options {
   multicast {
      scope-policy multicast-boundary;
   }
}
policy-options {
   policy-statement multicast-boundary {
      term bad-groups {
        from {
           route-filter 224.0.1.2/32 exact;
           ...
```

Cisco data plane scoping example

```
interface FastEthernet1/1
 ! must be applied to each interface
 ip multicast-boundary bad-groups
!
ip access-list standard bad-groups
 deny 224.0.1.2
 ...
```

Some appropriate IP protocols to multicast address space

• UDP[17]	various addresses (and ports*)
• IGMP[2]	various addresses, link local
• PIM[103]	224.0.0.13, link local
• OSPF[89]	224.0.0.[5|6], link local
• VRRP[112]	224.0.0.18, link local
• EIGRP[88]	224.0.0.10, link local
• ICMP[1]	?

TCP should be a no-brainer

- Though some systems are known to
  - return a SYN/ACK to a multicast SYN or
  - return a RST to lone ACK and
    - reply using the group addr as the src addr
- Poorly written worms may try scanning 224/4
- Edge networks should filter

ICMP to multicast addresses?

- Some systems respond to multicast echoes
  - e.g. IOS, Linux (RedHat), MacOS, Solaris
- RFC 2588 IP Multicast and Firewalls says:
  - “should only be used on the intranet”
- If you think directed broadcasts are a bad...
- Edge networks should probably filter

Luckily UDP scans don't work

RFC 1122 (Host Requirements) says:

“An ICMP error message MUST NOT be sent as the result of receiving [...] a datagram destined to an IP broadcast or IP multicast address”

Modern systems seem to behave properly.

Cisco protocol filtering example

```
interface FastEthernet1/1
 ip access-group interface-in in
!
ip access-list extended interface-in
 permit udp any 224.0.0.0 15.255.255.255
 permit igmp any 224.0.0.0 15.255.255.255
 permit pim any host 224.0.0.13
 permit ospf any host 224.0.0.5
 permit ospf any host 224.0.0.6
 deny ip any 224.0.0.0 15.255.255.255 [log]
 ...
```

How much multicast is enough?

- Do you want an ingress full of multicast traffic?
- Be careful not to starve control traffic
- Kind of like BGP max prefix – not for everyone
- 10-20% of link capacity?
  - unlimited for trusted high-speed sources?
- You could rate limit ingress control traffic too
  - we'll get to this in a bit

Cisco rate limiter example

```
class-map match-all udp-multicast
 match access-group udp-multicast
!
ip access-list extended udp-multicast
 permit udp any 224.0.0.0 15.255.255.255
!
! see next page
```

Cisco rate limiter example [continued]

```
policy-map edge-limiter
 class udp-multicast
  police 10000000 bc 5000 be 5000 \
  conform-action transmit exceed-action \
  drop
!
interface Fasthernet1/1
 service-policy input edge-limiter
```

IGMP

- Every router interface maintains IGMP state for each group any host has expressed interest in
- This in turn triggers PIM state
    - PIM DR maintains necessary interface state
        - usually at least two entries (*,G and S,G)
    - Upstream router(s) may have PIM state to get group traffic from sources to receivers

Cisco router IGMP mitigation example

```
interface FastEthernet1/1
 ! max groups
 ip igmp limit <1-64000>
 !
 ! IGMP join filter
 ip igmp access-group igmp-filter
!
ip access-list standard igmp-filter
 deny 224.0.1.2
 ...
```

IGMP snooping switches

- Switch inspects all packets for IGMP messages
    - learns router through IGMP queries
        - can a host masquerade as a router?
    - learns group membership from joins/leaves
- Switch builds port to group mappings
- Switch filters and spoofs IGMP joins/leaves

Cisco switch IGMP mitigation example

```
ip igmp profile 1
  range 224.0.1.2 224.0.1.2
!
interface FastEthernet1/1
 ! max joins permitted on this interface
 ip igmp max-groups [1-256]
 !
 ! IGMP filter list
 ip igmp filter 1
```

Cisco switch multicast mitigation example [continued]

```
interface FastEthernet1/1
 ! filter unknown multicast
 switchport block multicast
 !
 ! limit multicast input rate
 ! warning: also affects broadcast traffic
 storm-control multicast level bps 10m
```

PIM

- IGMP join results in (*,G) and (S,G) PIM state
  - per interface, plus upstream interface
  - plus state in upstream routers
- Mcast sender causes (*,G) and (S,G) PIM state
  - per interface, plus upstream interface
  - plus state in upstream routers and RP/MSDP
  - plus register, encap and decap effort

Cisco PIM filter examples
Mitigate PIM spoofing

```
interface FastEthernet1/1
 ! limit who your PIM neighbors are
 ip pim neighbor-filter pim-filter
 !
ip access-list standard pim-filter
 ! PIM neighbor router IP address
 permit 192.0.2.1
```

Cisco PIM filter examples

Limit groups the RP supports

```
! Setup RP and groups RP supports
ip pim rp-address 192.0.2.1 RP-ACL override
!
ip access-list standard RP-ACL
 deny 224.0.1.2
 ...
```

Cisco PIM filter examples
Limit register messages to RP

```
! set register pps rate limit
ip pim register-rate-limit 10
!
! filter remote sources
ip pim accept-register list registers
!
ip access-list extended registers
 deny ip any host 224.0.1.2
 ...
```

MSDP

- Each mcast sender/group pair is one MSDP SA
- Data may be included in an SA message
  - requires receiver MSDP peer to do work
- Misbehaving hosts cause SA floods to peers
  - akin to deaggregated BGP route leaks
  - except with additional processing workload
- Hardly anyone uses MD5 sigs in MSDP peering

MSDP SA limits

- Rate limiting a total number of SAs isn't enough
  - a single host can reach that limit easily
- You need global, per peer and per source limits
  - yes, it takes resources to manage this

Juniper MSDP SA filter example global limits

```
protocols {
    msdp {
        active-source-limit {
            maximum 26000;
            threshold 25000;
        }
    }
```

Juniper MSDP SA filter example peer limits

```
protocols {
    msdp {
        group MSDP-peers {
            peer 192.0.2.1 {
                active-source-limit {
                    maximum 15000;
                    threshold 14000;
                }
            }
        }
    }
```

Juniper MSDP SA filter example per-source limits

```
protocols {
    msdp {
        source 0.0.0.0/0 {
            active-source-limit {
                maximum 500;
                threshold 450;
            }
        }
    }
```

MBGP

- Routes with NLRI = 2 are multicast reachable
  - but really only assures ASN reachability
- You should be able to manage these routes like you do your unicast prefixes

For complete config examples

- Internet2 Multicast Cookbook
  - http://multicast.internet2.edu
- Juniper BCP in Multicast Security
  - http:www.juniper.net/solutions/literature/app_note/
- Secure Multicast Configuration Guide
  - http://aharp.ittns.northwestern.edu/papers/

A multicast minimalist approach

- Filter net ingress 224/4 at user edge interfaces
- Statically join groups you want on the interfaces
- Allow only those IGMP groups on switches
- Switches still manage joins, leaves and tables
- Presto change-o!
  - Relatively simple and safe, but limited receiver-only model - Good for video over IP

Could just do SSM, but...

- Widespread deployment yet to happen
- IGMPv3 does have one little *state* wrinkle
    - can specify multiple (*/S,G) per join
        - That is potentially a lot of additional state
        - increased risk of “IGMP join” flood attack

Edge problems still exist

- Hosts could still flood to group(s) in L2 domain
  - could use “private VLANs”
- Hosts could attempt to impersonate PIM DR
  - have real DR ignore edge PIM messages

A multicast minimalist approach configuration example

```
interface FastEthernet1/1
 ip igmp static-group 239.192.0.1
 ip access-group no-multicast-in in
!
ip access-list extended no-multicast-in
 deny ip any 224.0.0.0 15.255.255.255
 permit ip any any
```

Useful things to monitor

- IGMP, MBGP, MSDP and PIM state
- Join, leave, assert and register packet counts
- Weekly multicast routing/announcement report?
  - highlight bogus (SAs)
  - rank total SAs by source, ASN
  - rank top/bogus multicast flows

IGMP SNMP OIDs

- .1.3.6.1.3 - experimental branch
  - 59 - IGMP
    - 1.1.1.1.11 - total number of joins
    - 1.1.1.1.13 - total number of groups
    - 1.1.2.1.3 - groups and interfaces
    - 1.1.2.1.5 - time group's been active

mroute SNMP OIDs

- .1.3.6.1.3 - experimental branch
  - 60 - mroute
    - 1.1.2.1.4 - group, source and upstream
    - 1.1.4.1.5 - interface in octets
    - 1.1.4.1.6 - interface out octets

PIM SNMP OIDs

- .1.3.6.1.3 - experimental branch
    - 61 - PIM
        - 1.1.5.1.3 - group and RP entries
        - 1.1.5.1.5 – group and RP state uptime

MSDP SNMP OIDs

- .1.3.6.1.3 - experimental branch
  - 92 - MSDP
    - 1.1.3 - SA count
    - other OID usage seems to vary by vendor

Wanted? Router hacks

- PIM authentication (e.g. MD5 shared secret)
- Per source IGMP rate limit
- Per source maximum destination flow limiter
- Generic RTBH for group addresses
- no ip forwarding [IGMP | ...]
- no service mcast-echo-reply
- ip pim sparse-mode passive-interface

Sender-based attack mitigation

- Avoid taking down the RP
- Filter source from sending to group or 224/4
- Scope attacked group(s) at edge interface
- Disable PIM or PIM filter nearest the sender
- Change RP for attacked group
    - PIM RP blackhole?

Receiver-based attack mitigation

- Avoid taking down the RP
- Filter or limit IGMP from receiver
- Scope attacked group(s) on upstream interface
- Disable PIM or PIM filter nearest the receiver
- Change RP for receiver
  - PIM RP blackhole?

Inter-domain attack mitigation

- Avoid taking down the RP or MSDP peer
- Filter or limit SA source/originator
- Black hole MBGP route of source/receive net
- Scope attacked group(s) on upstream interface
- PIM filter nearest the receiver/sender
- Note: you have limited control outside your AS

Observation

- Unlike unicast attacks, you often don't have to rely on others to help you save your own net
- By limiting/filtering SAs you can reduce noise, state and processing workload
- By limiting/filtering joins/registers, you can reduce or eliminate certain multicast traffic from ever traversing your network

Multicast monitoring and tools

- Marshall Eubanks' multicast status page
  - http://www.multicasttech.com/status/
- UCSB Networking and Multimedia System Lab
  - http://www.nmsl.cs.ucsb.edu/
- NLANR/DAST Multicast Beacon
  - http://dast.nlanr.net/Projects/Beacon/
- Multicast router state summary script
  - http://aharp.ittns.northwestern.edu/software/

Recently expired, but relevant IETF drafts

- PIM-SM Multicast Routing Security Issues and Enhancements
  - draft-ietf-mboned-mroutesec-04
- Last-hop Threats to Protocol Independent Multicast (PIM)
  - draft-savola-pim-lasthop-threats-01
- IPv4 Multicast Unusable Group and Source Addresses
  - draft-ietf-mboned-ipv4-mcast-unusable-01

Some research-oriented papers

- Detection and Deflection of DoS Attacks Against Multicast Source Discovery Protocol
  - P. Rajvaidya, K. Ramachandran, K. Almeroth
- Deployment Issues for the IP Multicast Service and Architecture
  - C. Diot, B. N. Levine, B. Lyles, et al.
- Multicast-Specific Security Threats and Counter-Measures
  - T. Ballardie, J. Crowcroft (1995!)

General references

- Marshall Eubanks' multicast-related talks
  - http://www.multicasttech.com/papers/
- Internet2 resources and wg-multicast mail list
  - http://multicast.internet2.edu

The End

You have survived another multicast talk!