Chesapeake NETCRAFTSMEN

IP Multicast, Best Practices and Control

Peter J. Welcher

Introduction

Welcome back! The last two articles covered some basics of CiscoSecure ACS. While there's a lot more to be said on that topic, I'm trying to cover a broad variety of topics in these articles, not rewrite the massive Cisco documentation. So this month we'll shift to something I've been doing recently as consulting work.

The broad topic area is IP Multicast. Judging by session attendance (one sample) at Networkers, IP Multicast has become hot, apparently due to Service Provider interest in "triple play" (or per Cisco, "quadruple play"), specifically delivery of IP-based video services.

Our focus in this article will be on Best Practices and Control of IP Multicast. The next section will explain why one might feel the need to do this. Along the way we'll do a very quick review of IP multicast, and also list what I consider to be IP multicast Best Practices. They fit "control" in the sense of making your delivery of IP multicast as robust as possible.

Previous articles covered the basics of IP multicast fairly thoroughly. I have included links to them immediately below. They tell you how to get started, and a little bit about how IP multicast works.

Document	URL
Troubleshooting Too Much Multicast	http://www.netcraftsmen.net/welcher/papers/multicast-toomuch.html
IP Multicast and PIM Rendezvous Points	http://www.netcraftsmen.net/welcher/papers/multicast04.html
PIM Sparse Mode	http://www.netcraftsmen.net/welcher/papers/multicast03.html
PIM Dense Mode	http://www.netcraftsmen.net/welcher/papers/multicast02.html
The Protocols of IP Multicast	http://www.netcraftsmen.net/welcher/papers/multicast01.html

Popular IP Multicast Applications

The usual question at this point is, why would I want IP multicast in my network? Some answers:

- Cisco (or other) Music on Hold for IP phones
- Nortel Symposium web, tracking Call Center skill groups and queue depths, mean wait times, etc.
- PC desktop imaging: multicast Norton ghost for example
- Tibco and PGM-based market data (brokers in various markets)
- Video "channels" the old IP/TV, or using the multicast transmission capability now in Cisco ACNS
- Novell or other user of IP Service Location Protocol (SLP)
- Microsoft Univeral Plug and Play (UPnP)

Surprises when deploying IP multicast:

- Multicast ghost traffic may quickly show up (local imaging subscribing to a group and hooking up with a ghost server elsewhere)
- UPnP is out there, massive numbers of sources
- People may have been using IP multicast locally, within a VLAN or subnet. When you enable it globally, these users find each other. For example, computer gaming.

Conclusion: turn IP multicast on gradually, and monitor it. You don't want to turn it on all over and then find massive new state in routers or massive new traffic flows on the network.

Multicast lesson that keeps getting re-learned (ghost, PGM): if you're distributing images, files, etc., then if there is a disparity in speeds such as 10 Mbps versus 100 Mbps, things generally don't work well. Use different multicast groups for the different speeds and things may work a lot better.

Why Control IP Multicast?

A quick review of IP multicast may help:

- IP multicast packets are sent to Class D multicast addresses, in the 224.0.0.0 - 239.255.255.255 range. These destination addresses are referred to as multicast groups.

- Receivers join a group to indicate interest in receiving traffic destined to that group. Routers then conspire to get traffic from sources to receivers. A newer approach called Source Specific Multicast (SSM) has receivers indicate which destination group and which source they wish to receive.
- When getting started, a source just starts transmitting. The adjacent router(s) determine whether forwarding is needed or not.
- When doing Protocol Independent Multicast (PIM) Sparse Mode, the source traffic is registered with the Rendezvous Point (RP), which is used to connect sources and receivers up. Receiver-adjacent routers then trigger creation of a source-specific tree for multicast propagation and replication. The RP is not needed when doing SSM.

IP multicast (IPmc) is still evolving somewhat, as Cisco and users learn about new requirements. There are a number of new technologies, to fill new needs, for example Source Specific Multicast. The old methods are not all going away, because there are different needs and situations in IP multicast.

We've seen 6500 switches reboot when rapid joins from too many (5000 or more) multicast sources filled the TCAM. After that, controlling (S, G) state made it onto my mental list of "things to control or look out for". I talked to some Cisco folks about this at 2003 Networkers. I now notice a bit more emphasis on state control in presentations.

Networkers 2006 this year contained talk RST-2262, *IP Multicast Security*. This appears to be a new theme for Networkers. I took this presentation as confirming that the need exists, and reinforcing the intent I had prior to that of writing this article. That presentation has a lot of good material in it -- high recommended! Some of does seem to be more Service Provider or very large organization focus, e.g. scoping and multicast boundaries. The presentation also talked about Firewalls (PIX 7.0) and IPmc, and encrypted IPmc, which are security-related but outside the scope I intend for this article. In this article, I'm trying for that medium level of control that protects your network to a "reasonable" level, without creating an adminstrative nightmare. If you're a Service Provider or otherwise need to lock down your network more tightly, then see the above Networkers presentation.

My second concern is admission control. What multicast traffic do we want on the network? I've seen several large networks with a lot of unknown multicast traffic on them. This worries me. Real-time multicast traffic can entail multiple large flows of data. Large unicast flows tend to be fairly isolated -- someone doing a file download here or there. Multicast can have broader impact.

This concern may not be completely logical: I can make a case that unicast has many of the same characteristics. What about:

- Someone widely deploys a new unicast-based application that starts pulling down huge amounts of data
- Enterprise anti-virus or patch control software triggers ongoing downloads of a patch (I've heard of this taking down a network twice now)
- People do peer-to-peer downloads of TV shows?

The difference is that generally unicast surprises require new application deployment, or many people doing something network-hostile with an existing application. Multicast surprises can happen if there are many sources, or if there are existing listeners for a group and then some source starts transmitting a large stream to the same group.

My current perspective is that multicast applications are generally new to vendors, and both vendors and many networking staff are still gaining experience with IP multicast. This seems like a situation where exercising a bit more control up front can reduce operational surprises. It is also a situation where IP multicast applications are not as prevalent as unicast ones, so there is a better chance of exerting some positive control.

Ok, so I'm a control freak! I do want to control unplanned IPmc applications.

I also want to keep rogue sources from using destination IPmc addresses that are legitimate, especially to avoid any kind of Denial of Service (DoS). Multicast also makes Distributed DoS somewhat easier -- just get a bunch of PC's sending multicast, it'll go to anywhere the destination group has been joined. If receivers for that group are widely distributed, then they might be spraying traffic to a lot of the network. The traffic would be sender-initiated, rather than receiver-initiated.

Much of this can be managed with access lists (ACLs) on all user interfaces, but that might not be very manageable. We discuss another partial alternative below.

Controls

The Cisco Networkers presentation already mentioned had slides talking about threats. Specifically:

- Attacks from sources to hosts (groups, in IPmc)
- Attacks from sources to networks
- Attacks from receivers (unauthorized content, bandwidth, router/switch)

The presentation goes on to examine other threats, control-plane security, and how to protect against router state overload (either (S, G) mroute entries or IGMP groups). It goes into attacks on IPmc protocols (PIM, IGMP, MSDP). And it also looks at controlling what groups receivers can join. All good stuff!

Practical example: one reason a provider might want to control IGMP is to ensure you only get the channels of digital TV appropriate for the package you're paying for. Is this like the settings controlling which cable TV channels your house receives?

Our focus here will be mostly the first two, and more in the context of inadvertent attacks than deliberate attacks. So we'll go on to look at:

- Controlling sources and multicast flows
- Preventing IP multicast source state in routers and switches from becoming a problem.

IP Multicast Design Best Practices

It seems helpful at this point to review IP multicast design Best Practices. If you're serious about doing multicast, you'll want a stable deployment. And you don't need to spend your time hunting down obscure problems. The standard reference

(IPmc System Reference Network Design Guide, or IPmc SRND Guide) is listed below. We'll look briefly at the conclusions I've reached, in essence a very short summary of that document.

The first thing to do is to either turn on IP multicast on all active interfaces interconnecting routers, or to learn how RPF checks work and plan carefully. The key point is that multicast arriving on a non-RPFinterface generally gets ignored. So your unicast routes back to Rendezvous Point (RP) or to the multicast source need to be on interfaces where multicast was enabled (at both ends). This can be particularly disconcerting if you had multicast routing, then a failure, and the alternative unicast routing path isn't multicast-enabled.

I highly recommend looking into multicast addressing and coming up with an address plan. It sure beats going back and doing things over again! Cisco has a great writeup on the topic, publically available. A link is included in the Reference Links table below.

Little-known fact: you want to avoid multicast addresses of the form <something>.0.0.x and <something>.0.1.x, also <something>.128.0.x and <something>.128.1.x. IGMP Snooping and its predecessor CGMP do a fine job of delivering an IPmc group to only the user ports that want it. But the multicast groups 224.0.0.x and 224.0.1.x are used to contact all multicast hosts and routers, also for routing protocols, so they have to go to all switch ports. The switch tracks this by destination MAC address. In terms of corresponding MAC addresses, unfortunately, 32 IP multicast groups map to the same MAC. The mapping ignores the first octet and the 128 bit in the 2nd octet of the destination multicast IP, keeping the last 23 bits of the IP address in the multicast MAC address. So not only do 224.0.0.x and 224.0.1.x get flooded to all ports in a VLAN, but so do 225.0.0.x, 225.0.1.x, 225.128.0.x, 225.128.1.x, 226.0.0.x, etc. If you're using one of these addresses, it is not the end of the world, but it may well be causing extra multicast at the edges of your network, and the CPU interrupts may be mildly slowing down affected PCs.

Your really should use PIM Sparse Mode (PIM-SM). PIM Dense Mode floods every 3 minutes, which is really disruptive if you have big multicast flows. You can either make sure you have a very robust Rendezvous Point, or only allow PIM-SM on interfaces. If you do the latter, PIM-DM fallback can still occur -- there is a recent command to prevent that. I generally want to use `ip pim sparse-dense-mode` on interfaces so that I can easily do AutoRP (next).

I generally prefer to use AutoRP. I also prefer to do Anycast RP, with two well-placed RPs talking MSDP to each other. These are generally in the Data Center, on well-connected high-capacity devices. I also do RP of Last Resort. For details, see the already-mentioned Cisco SRND document.

For state control, see the next section.

I generally enable standard IPmc, Any Source Multicast ("ASM"). I would use Bidirectional PIM except that (a) many sites are not running new enough code yet, and (b) using Bidir PIM precludes a nice option that we'll cover in the next section. The Cisco presentation mentions SSM. I like the idea of reserving an address range for SSM. But SSM is new, support for it in host still limited, and SSM creates source-specific multicast trees, which I've become highly allergic to. It seems like SSM solves one security problem, and risks another. Or am I just resisting a new idea?

Let's not worry about IPv6 multicast in this article. It's mostly the same, with some differences.

IP Multicast and Access Lists

You can use extended access lists to block multicast packets. The Cisco recommendation is to stick with protocol "ip" for this. You cannot block the special ranges of local multicast mentioned above. Address 0.0.0.0 matches (*, G) traffic flows but does not block (S, G) traffic (practical use?).

So if your servers are in address block 10.1.1.0 /24, you might only allow multicast coming from official servers. Create the following access list (ACL) and apply it to all inbound interfaces.

```
ip access-list extended ipmc-source
  permit ip host 10.1.1.0.0  0.0.0.255  224.0.0.0  15.255.255.255
  permit ip any 224.0.0.0  0.0.1.255
  deny ip any 224.0.0.0  15.255.255.255  log
  permit ip any any

interface ethernet0
  ip access-group ipmc-source in
```

There's a pitfall here that I have seen in slideware, and almost fallen into myself. You start thinking multicast, and you can easily forget the later permit statements at the end of the above ACL. Why are those permit statements necessary? How about unicast server traffic? Local routing and control multicasts (EIGRP, OSPF hellos)?

You could of course tighten this down. If you're using part of 239.0.0.0 /8, then restrict to that block instead of the whole 224 block, as far as destinations.

The main advantage of this approach is simplicity and clarity. The drawback is applying an ACL to any inbound interface a multicast source might be on -- all user or server subnets. Maintenance isn't too bad, as long as you keep the same ACL on all devices. Use CiscoWorks or your favorite tool to push out updates.

You can also use ACLs to control allowed receivers and the groups they join.

```
ip access-list extended allowed-multicast-igmp
  permit ip any host  239.2.3.4
  permit ip 10.0.0.0  0.255.255.255  239.3.0.0  0.0.255.255
  deny ip any any

interface ethernet 0
  ip igmp access-group allowed-multicast-igmp
```

Note the ACL is applied with a slightly different command, the "ip igmp access-group" command. This allows IGMP from any host to join group 239.2.3.4, and from any host in network 10 /8 to join groups in the 239.3.x.x range. Unless you're doing different things for different IP addresses (not likely with DHCP), you can probably use source 'any' on a per-VLAN

(per-subnet) basis, and just specify the multicast groups hosts are allowed to join.

Caution: Networkers slides say any deny statements must be protocol "ip" or they won't be effective.

How to Control IP Multicast

The "**`ip pim accept-register`**" command might be useful to you. You put it on your RP and it exerts remote control. Specifically, when a source fires up multicast, the adjacent router sends a Register packet to the PIM-SM RP. The RP then checks the ACL specified in this command. If the source and group pass the ACL, the Register process is completed. Otherwise, the RP sends back a Register-Stop message. This blocks that multicast flow from ever getting forwarded beyond the directly adjacent router. You will see (S, G) state in that one router, since the router needs some way to record the Register-Stop information.

From careful lab observation, there is a gotcha to this. Local receivers that sent in an IGMP join do get forwarded the blocked multicast stream, even if on a different local VLAN (subnet). This was not what I really wanted to have happen. It makes logical sense, and does protect network bandwidth. But if you are looking to only maintain multicast source access lists on the RP, you'll be as disappointed as I was. Your choice: use this command to mostly control multicast, or lock down multicast with edge ACLs.

Here's what that might look like in practice. We re-use our ipmc-source ACL from above. On the two Anycast RPs, we configure:

```
ip pim accept-register list ipmc-source
```

By the way, you cannot use this command with Bidirectional PIM. That puts it slightly at odds with the objective of state reduction (or is "containment" a better word for this?).

Coming Cisco IOS code will apparently allow controlling all use of multicast groups with an ACL, via the following command:

```
ip multicast group-range <standard ACL>
```

The 12.4 T code has such a feature for IPv6 already.

To minimize (S, G) state with ASM (Any Source Multicast), we can use the SPT threshold command. This needs to be put into every router running PIM:

```
ip pim spt-threshold infinity
```

Note that using the command as shown has the consequence that NO multicast groups will use the SPT source-specific tree. This is fine for campus or well-connected RPs. It might not be so good for remote RPs where there is a better path between source and receiver. That's the trade-off: optimal paths for each source and multicast group, versus less state in the router. Bidir PIM entails the same trade-off.

For more flexibility, the above command can be entered with an access list, to specify which groups it applies to. The same applies to Bidir-PIM.

To prevent a bogus PIM message for a bogus RP, you can use the command

```
ip pim accept-rp auto-rp
```

If you prefer, you can specify precisely which addresses can act as RP for join and prune messages by substituting an ACL name for "auto-rp" in the above.

For control plane state limiting, there are now all sorts of limit commands, to limit MSDP-learned sources, IGMP entries, and mroute state. Many of these can be modified with an ACL to limit state by blocks of multicast groups. You can even do cost-based limits, as a form of Call Admission Control. Some of these also allow a syslog warning threshold, similar to some of the Service Provider features one now has available to limit VRF routing table size growth in MPLS VPNs.

Syntax examples:

```
ip multicast route-limit <limit> [ <threshold> ]
ip msdp sa-limit <peer> <limit>
ip igmp limit <n> [ except <ext-acl> ]
ip multicast limit [ rpf | out | connected ] <ext-acl> <max>
ip multicast limit cost <ext-acl> <multiplier>
```

You can filter out undesirable traffic for the control plane with the following command:

```
ip receive access-list <ext-acl>
```

If older code doesn't support this, you can use per-interface ACLs as above to do the same thing.

Another thing you can do for control plane protection is use COPP, control plane policing, in chassis that support it. The idea is to police any traffic to a multicast destination that are received by the control plane in the router, to some tolerable bit per second rate.

Summary

A word about multicast. You may see surprise entries in "**`show ip mroute`**". I've been noticing them a lot lately, because the **`show ip mroute`** command does show table state, and some entries are needed to track that state, for example multicast replication path from source to Rendezvous Point.

If you have not yet downloaded and read them, I strongly encourage you to look at the Cisco *Guidelines for IP Multicast Address Allocation* and *IP Multicast SRND* documents. Links are in the following reference links table.

Here are some of the best reference links relating to IP multicast:

Document	URL
Cisco IP Multicast technology page	http://www.cisco.com/en/US/products/ps6552/products_ios_technology_home.html http://www.cisco.com/go/multicast
IP Multicast White Papers	http://www.cisco.com/en/US/tech/tk828/tech_white_papers_list.html http://www.cisco.com/en/US/products/ps6552/prod_white_papers_list.html
Guidelines for Enterprise IP Multicast Address Allocation	http://www.cisco.com/en/US/tech/tk828/technologies_white_paper09186a00802d4643.shtml
Cisco AVVID Network Infrastructure IP Multicast Design (SRND)	http://www.cisco.com/application/pdf/en/us/guest/tech/tk363/c1501/ccmigration_09186a008015e7cc.pdf
Cisco "Deploying IP Multicast" presentation, Networkers 2006	http://www.networkersonline.net/files/subsystems/12073/RST-2261(USA,2006).pdf Note: Log into Networkers Online first. You'll then be able to follow this link. Note2: To get a login, you need to subscribe, or to have been a paid attendee at Networkers. See http://www.networkersonline.net/ for a free sample session and for the Subscription link.
Cisco "IP Multicast Security" presentation, Networkers 2006	http://www.networkersonline.net/files/subsystems/12073/RST-2262(USA,2006).pdf Note: Log into Networkers Online first. You'll then be able to follow this link. Note2: To get a login, you need to subscribe, or to have been a paid attendee at Networkers. See http://www.networkersonline.net/ for a free sample session and for the Subscription link.

I hope you enjoyed this article.

Your comments, questions, and suggestions for future articles are of course welcome! See below to decipher my email address.

Dr. Peter J. Welcher (CCIE #1773, CCSI #94014, CCIP) is a Senior Consultant with Chesapeake NetCraftsmen. NetCraftsmen is a high-end consulting firm and Cisco Premier Partner with multiple specializations, dedicated to quality consulting and knowledge transfer. NetCraftsmen has nine CCIE's, with expertise including large network high-availability routing/switching and design, IP Telephony, QoS, MPLS, IPSec VPN, wireless LAN and bridging, network management, security, IP multicast, and other areas. See http://www.netcraftsmen.net for more information about NetCraftsmen. Pete's links start at http://www.netcraftsmen.net/welcher . New articles will be posted under the Articles link. Questions, suggestions for articles, etc. can be sent to **pjw <at> netcraftsmen <dot> net**.

8/7/2006