SCTE CABLE-TEC EXPO'13
OCTOBER 21-24 / ATLANTA, GA

SCTE Society of Cable Telecommunications Engineers

Routing the Cable Network: Unicast, Multicast and MPLS Protocols

An Overview of Common Network Routing Protocols & Their Intersections with Cable Architectures

A Technical Paper prepared for the Society of Cable Telecommunications Engineers
By

Jay Herbert
Systems Engineer
Cisco Systems Inc.
5030 Sugarloaf Parkway, Lawrenceville, GA 30044
678.352.2763
jherbert@cisco.com

Overview

This paper provides a two-part assessment of how the network routing protocols of the Internet -- and specifically unicast, multicast and MPLS-based routing protocols --are used in today's cable broadband architectures. Additionally, the paper analyzes routing information and use cases gathered by the author from engineering peers and colleagues within the top 12 North American MSOs. An analysis of the MSO data is provided, to identify trending and popularity (or lack thereof) of certain routing protocols.

The goal of this paper is to impart a basic level of understanding about how routing protocols of the Internet (north of the Cable Modem Termination System, /CMTS) are used within cable architectures. Network routing protocols continue to evolve, but they all have a common starting point and use case -- an origin point from which all change takes place.

In cable, that starting point is in the transition of legacy (QAM-based) services to all-IP. After all: Cable Modem Termination Systems are, in part, routers. The Internet uses routers. Lots of them. The trickle-down of Internet protocols into Cable Access Plant is inevitable, and it’s ultimately a good thing.

But in order to understand the more complex use cases, one must first understand the basics. That’s Part 1: Why routing protocols exist, and how they work. Part 2 examines how the different network protocols are used inside a cable’s network architecture. Informed by the results of a query of a dozen North American MSOs, the analysis ascertains the industry’s current and strategic direction with respect to networking protocols.

PART I: NETWORK PROTOCOL BASICS

As network engineers and technical leaders, we use several variables in order to assist in the architectural decision making process. One variable relied upon is what our peers are doing. How do cable service providers use routing protocols? Can trends be established by analyzing this data?

For this paper, we gathered the current routing protocol architecture and future plans of the top 12 North American MSOs. The analysis of that data shows some clear trends regarding the usage -- and avoidance -- of certain routing protocols. This data highlights what architectures are popular, and what forwarding mechanisms still remain the most common.

"If you are not riding the wave of change, you'll find yourself beneath it." This common motivational poster quote could apply to almost any aspect of life. It certainly applies to networking. Our world of networking is constantly changing. Nothing stays static for long except maybe that one lone router everyone seems to have in the network with an uptime of greater than 10 years.

For instance, it is helpful to understand the fundamental basics of Border Gateway Protocol and what a cable provider might want to accomplish with it. BGP is used among service providers to essentially tell each other about the networks (subnets) they own. In this way, when a router within a service provider network receives a packet, it knows to whom it belongs. Maybe it belongs to a local customer, another service provider across the country, or maybe to a network from the other side of the globe.

This paper will explore the many aspects of routing protocols and how they are used within cable. Although detailed, it is not an exhaustive analysis of how routing protocols work, nor is it a complete assessment of their multiple use cases. How they are used in most cable architectures will be the primary focus.

THE PURPOSE OF ROUTING PROTOCOLS IN CABLE

A routing protocol is used to distribute information that routers use to make forwarding decisions. This forwarding decision might be based on an IP address (e.g. destination-based forwarding) or an abstraction, such as a label (e.g. Multiprotocol Label Switching, or MPLS.) Forwarding decisions can also be based on pre-established, “connection-oriented” forwarding paths or demand-based trees, which are set up via a routing protocol process. Two examples we will discuss in this paper are Traffic Engineering and Multicast forwarding.

The distribution of forwarding information is the main role of a dynamic routing protocol. They have other specialized functions and features, such as layer convergence, route selection, scale and policy control, among others, but the main purpose is to share information dynamically. If a dynamic routing protocol is not used, then the static configuration of each router along a path from source to destination is required in order to forward a packet. Although static routing is used for small solutions, it is not a practical large network routing solution. In cable and almost all networks we see static routing used to control and steer traffic locally, in support of other features.

Routing protocols differ in how they accomplish the task of information distribution. Just like there are different tools in your garage to accomplish different tasks, certain routing protocols are used for specific purposes.

Interior Gateway Protocols or IGPs are used to distribute routing information within an autonomous system or single operational entity. They were designed to be fast and efficient, with minimal policy enforcement capabilities. Examples of common IGPs used in cable are OSPF, ISIS and RIPv2.

An Exterior Gateway Protocol, or EGP, is a high level designation that has since been replaced with Border Gateway Protocol (BGP), largely because BGP has been established as the only acceptable standard for an EGP. BGP is used to share routing information between autonomous systems. It is used to distribute information between companies and hence has a high level of policy enforcement capabilities and scale. All Internet routes are globally distributed using BGP.

Multicast routing protocols are used to create distribution trees that are typically (but not always) based on a demand model. If you seek this-or-that multicast service, you ask for it (create a demand) and the multicast routing protocol will build a forwarding path to you.

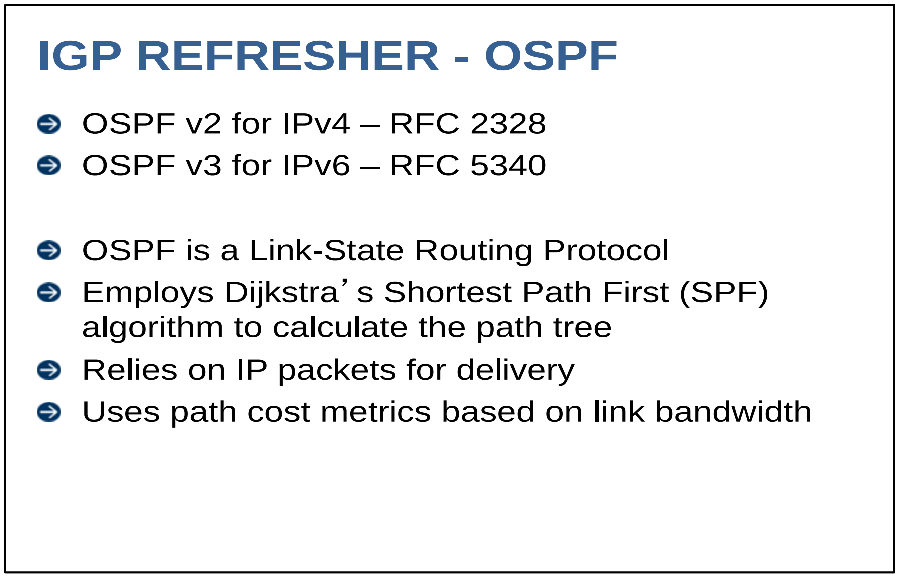
Multiprotocol Label Switching or MPLS is a routing protocol that assigns labels to routing information and distributes those labels. Those labels become the mechanism to make forwarding decisions, and are typically derived from the knowledge of other routing protocols. Because this level of abstraction can be layered (i.e. label stack), the ability to make complex forwarding decisions is created. This has created the ability for MPLS to become a very flexible service creating protocol. MPLS is used today to create different routing topologies for different groups or VPNs (Virtual Private Networks.) MPLS also has the ability to carve out (establish) a path used for Layer 2 activities, like Pseudowire services (a logical connection between two end points.) Multicast routing protocols and MPLS actually use the topology awareness features of other routing protocols.

REFRESHER: OSPF, ISIS & RIP

The high level overview of Interior Gateway Protocols will focus specifically on three of them, because of their current usage within cable architectures:

- RIP (Routing Information Protocol),
- OSPF (Open Shortest Path First) and
- ISIS (Intermediate System to Intermediate System).

IGPs such as IGRP (Interior Gateway Routing Protocol) and EIGRP (Enhanced Interior Gateway Routing Protocol) are not included here but are mentioned briefly within the use case portion of this paper.



IGP REFRESHER - OSPF

- OSPF v2 for IPv4 – RFC 2328
- OSPF v3 for IPv6 – RFC 5340

- OSPF is a Link-State Routing Protocol
- Employs Dijkstra's Shortest Path First (SPF) algorithm to calculate the path tree
- Relies on IP packets for delivery
- Uses path cost metrics based on link bandwidth

Figure 1 Basic characteristics of Open Shortest Path First protocol

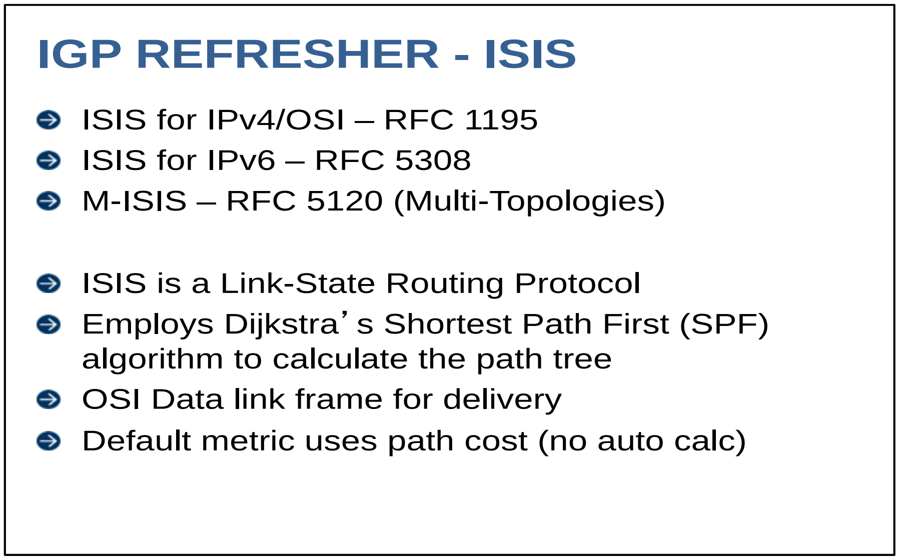
OSPF stands for "Open Shortest Path First" (and is one of the rare acronyms that contains an actual verb.) Two versions of OSPF are used in today's cable networks: v2 and v3. The difference between these versions is their ability to work with either IPv4 (using OSPFv2) or IPv6 (using OSPFv3) routes.

OSPF is a link state routing protocol. That means that subnet reachability is distributed to everyone within an area, until all participating nodes have identical databases (network knowledge). The Shortest Path First algorithm is then run independently on each router. OSPF uses Internet Protocol for router updates, like Link State Advertisements, or LSAs. Note: You will soon see this is not the case for all IGPs. The metric used to determine the best path is the lowest transit cost of available bandwidth calculated between source and destination. Transit cost is based on a link bandwidth calculation. OSPF uses a 2-layer model, where each non-zero area is connected to the backbone or Area 0. The boundary between areas is within the routers themselves, since a router's links can be in different areas.

There are some important "Types" in OSPF (which also go by "Router Types"), which determines the role a router plays. Internal, ABR (Area Border Router) and ASBR (Autonomous System Boundary Router) are the common types. They are defined by their position in the network and what areas they touch. There are also several different types of LSAs (Link State Advertisements), depending on the information being sent from the router to other routers. The type of area created can control which LSAs are allowed across the boundary between areas.

Similar to OSPF, ISIS (pronounced as the constituent letters and an acronym for Intermediate System to Intermediate System) is a link state routing protocol that uses a two level hierarchy. In ISIS, the boundary between areas is the link itself and not the router, like it is with OSPF. ISIS routers are designated as being Level 1 (intra-area), Level 2 (inter-area) or Level 1-2 (both). Routing information is exchanged between Level 1 routers and other Level 1 routers. Likewise, Level 2 routers only exchange information with other Level 2 routers. It stands to reason that Level 1-2 routers exchange information with both levels and are used to connect the inter-area routers with the intra-area routers.

A significant difference between ISIS and OSPF is that the "hello intervals" and hold times between two ISIS neighbors do not have to match. Each router honors the hold time advertised by its neighbor. Another interesting difference between OSPF and ISIS is their adjacencies. ISIS considers routers adjacent as soon as they exchange Hellos.



IGP REFRESHER - ISIS

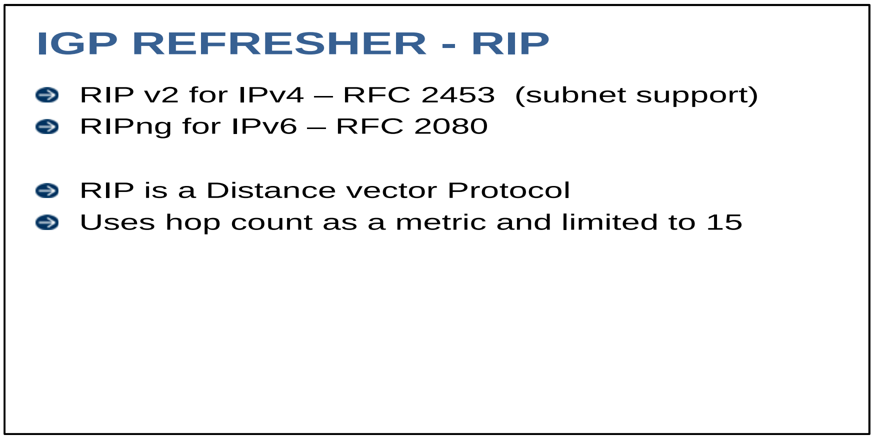
- ISIS for IPv4/OSI – RFC 1195
- ISIS for IPv6 – RFC 5308
- M-ISIS – RFC 5120 (Multi-Topologies)

- ISIS is a Link-State Routing Protocol
- Employs Dijkstra's Shortest Path First (SPF) algorithm to calculate the path tree
- OSI Data link frame for delivery
- Default metric uses path cost (no auto calc)

Figure 2 Basic characteristics of Intermediate System to Intermediate System Protocol

ISIS allows three types of routing domains: OSI (Open Systems Interconnect), IP (Internet Protocol) and DUAL (Diffusing Update Algorithm.) OSI domains do not exist anymore, at least in cable networks. Extensions to ISIS (such as RFC 5308) have been added to support IPv6.

Multi-topology ISIS allows the user to carve out two separate topologies. Although not mandated, the purpose is typically for the support of two distinct IPv4 and IPv6 topologies. Multi-topology allows you to run a separate SPF per topology. ISIS has 4 metrics, but only one is required. The default of cost is typically used. Unlike OSPF, the cost is not calculated based on link speed. Rather, the network designer bases it on an arbitrary number set.



IGP REFRESHER - RIP

- RIP v2 for IPv4 – RFC 2453 (subnet support)
- RIPng for IPv6 – RFC 2080

- RIP is a Distance vector Protocol
- Uses hop count as a metric and limited to 15

Figure 3 Basic characteristics of Routing Information Protocol

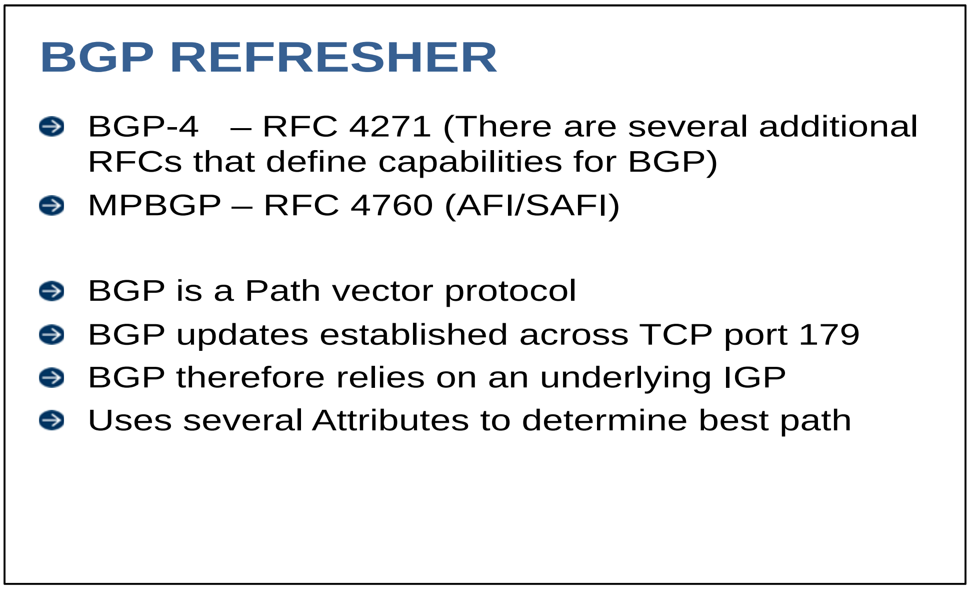
There are currently three types of Routing Information Protocol (pronounced like the word “rip”): The original version, known as RIPv1; RIPv2, which added subnet support; and RIPng (where the “ng” stands for “next generation”), which supports IPv6.

RIP is a distance vector protocol and therefore it only sends its neighbor the reachability information that it knows about. A big limitation of RIP is its limited hop count of 15. This means you cannot run RIP on a network that is longer then 15 hops from any source to destination. For those scale-related reasons, RIP is not widely deployed.

REFRESHER: BGP

Today, the only accepted standard for an Exterior Gateway Protocol (EGP) is Border Gateway Protocol (BGP). All service providers use BGP to advertise their Internet routes to other service providers.

There are several Requests for Comments (RFCs) within the Internet Engineering Task Force (IETF), which define extensions to BGP. Two are included here as good starting points.



BGP REFRESHER

- BGP-4 – RFC 4271 (There are several additional RFCs that define capabilities for BGP)
- MPBGP – RFC 4760 (AFI/SAFI)

- BGP is a Path vector protocol
- BGP updates established across TCP port 179
- BGP therefore relies on an underlying IGP
- Uses several Attributes to determine best path

Figure 4 Basic characteristics of Border Gateway Protocol

BGP is a path vector protocol, or a variant of a distance vector routing protocol. It maintains path information to avoid loops.

A common companion term used in BGP discussions is NLRI, or Network Layer Reachability Information. This is the information in BGP messages that identifies reachability within BGP update messages.

Two other very common terms are iBGP, or internal BGP, and eBGP, or external BGP. These terms are used to identify the type of BGP peer being established between routers. The difference between these two types will depend on the autonomous system (AS) of each peer. Two peers within the same AS are considered to have an iBGP connection. Two BGP peers from different autonomous systems are considered to have an eBGP connection. Each connection type follows different rules within the BGP protocol.

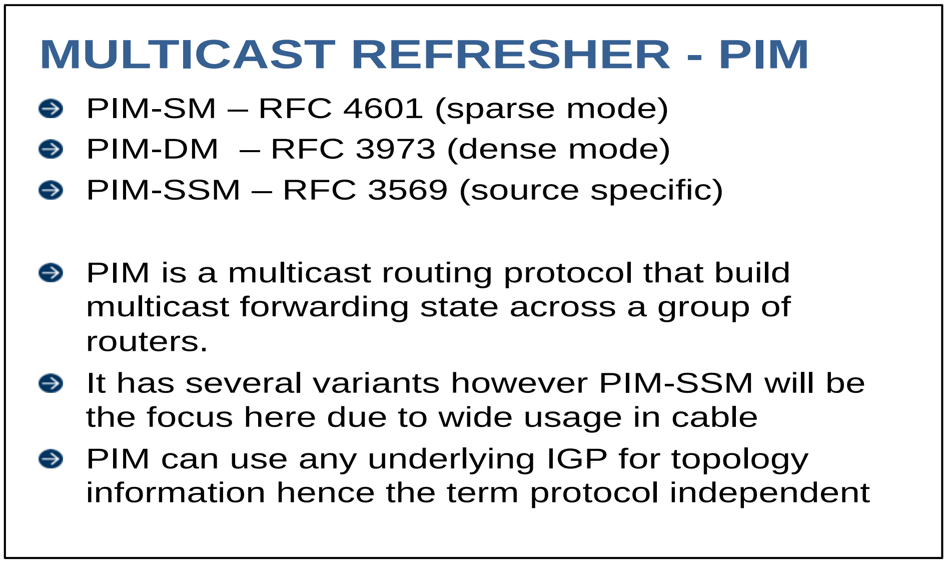
Multiprotocol BGP, or MPBGP, added the ability to advertise more information then just unicast ipv4 reachability. Some common terms used within MPBGP are Address Family Identifier (AFI) and the Subsequent Address Family Identifier (SAFI). These fields, in combination, identify the set of network layer protocols to which the address carried in the “next hop” field must belong.

An important term and concept within BGP is the iBGP full mesh requirement. This rule requires all BGP Provider Edge (“PE”) routers within the same iBGP peer to have established peering connections. We will see later in the paper how most cable operators solve this requirement. BGP attribute modification, policy control and route selection are all important concepts in BGP, and explain why it is so widely accepted as the Internet standard routing protocol. Part 2 of the paper examines cable network BGP use cases.

REFRESHER: THE MULTICAST PROTOCOLS

The multicast families of protocols are used to build “trees” or reachability between multicast listeners and multicast sources. Native Multicast forwarding is accomplished with primarily two protocols. The first is Protocol independent Multicast (PIM) so named because it can use any underlying IGP to build multicast trees.

The second primary multicast protocol is Internet Group Management Protocol (IGMP). IGMP is a client protocol used to dynamically signal the first hop router of its interest in receiving a particular multicast group.



MULTICAST REFRESHER - PIM

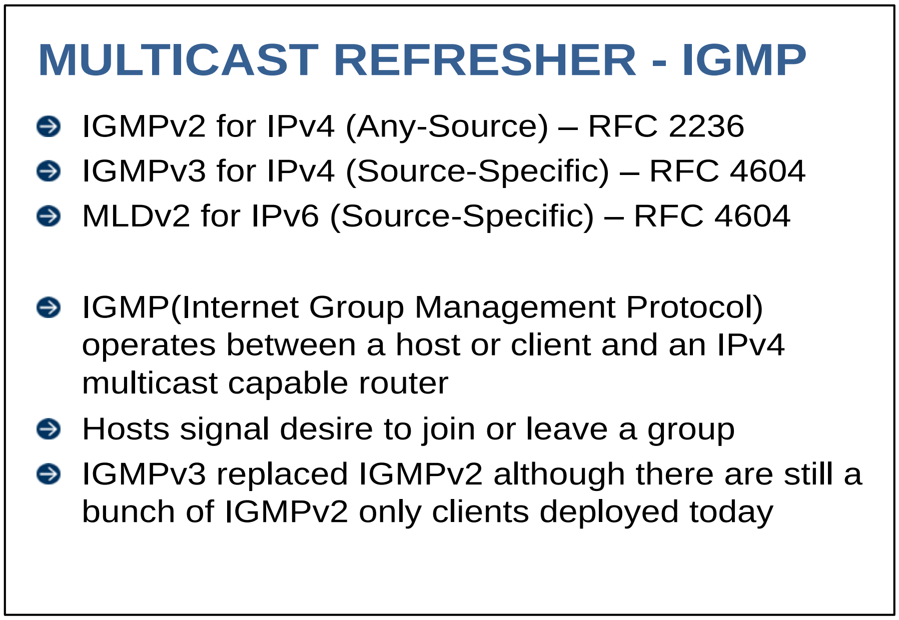
- PIM-SM – RFC 4601 (sparse mode)
- PIM-DM – RFC 3973 (dense mode)
- PIM-SSM – RFC 3569 (source specific)

- PIM is a multicast routing protocol that build multicast forwarding state across a group of routers.
- It has several variants however PIM-SSM will be the focus here due to wide usage in cable
- PIM can use any underlying IGP for topology information hence the term protocol independent

Figure 5 Basic Characteristics of Protocol Independent Multicast

PIM is a multicast routing protocol that builds multicast forwarding state across a group of routers. It has several variants, however, the focus in this paper is on PIM-SSM, or Source Specific Multicast, because it is the most popular variant deployed in today’s cable networks.

PIM-SSM means that multicast forwarding state is built within a router, based on both the source and destination addresses. It’s written out as (S,G) (pronounced “S comma G”), where S stands for Source IP address and G stands for Group IP address.



MULTICAST REFRESHER - IGMP

- IGMPv2 for IPv4 (Any-Source) – RFC 2236
- IGMPv3 for IPv4 (Source-Specific) – RFC 4604
- MLDv2 for IPv6 (Source-Specific) – RFC 4604

- IGMP(Internet Group Management Protocol) operates between a host or client and an IPv4 multicast capable router
- Hosts signal desire to join or leave a group
- IGMPv3 replaced IGMPv2 although there are still a bunch of IGMPv2 only clients deployed today

Figure 6 Basic characteristics of Internet Group Management Protocol

IGMP allows a client to advertise its interest in a particular multicast group. IGMPv3, in particular, provides a dynamic way for clients to advertise their interest, by including the source as well as the group address. By contrast, the original IGMPv2 only allowed clients to signal interest in the group address, not the source. This forced the routing network to determine who the source was, which significantly increased the complexity of multicast routing. Today, IGMPv3 (when available) is the desired protocol.

Multicast Listener Discovery, or MLD, is used by IPv6 routers and is similar in concept to IGMP for IPv4. The protocol is embedded in ICMPv6 (Internet Control Messaging Protocol), instead of using a separate protocol.

MLDv1 is similar to IGMPv2, and MLDv2 is similar to IGMPv3.   There are no deployments of multicast IPv6 in any cable network as of the writing of this paper and therefore no use cases will be discussed.

MPLS

MPLS uses labels to make forwarding decisions vs. the IP destination field in an IP packet. MPLS uses many of the underlying packet-based routing protocols in order to function. Its popularity is due to the complexity of services it can enable.

There are several control protocol options when using an MPLS forwarding architecture. The main idea is to associate some type of FEC (Forward Equivalence Class) to a set of labels. Examples of protocols included in an MPLS-architected network include Label Distribution Protocol (LDP), Resource Reservation Protocol (RSVP), Constraint-based Shorted Path First (CSPF), Multi-protocol Border Gateway Protocol (MPBGP), or a combination of all of them.

MPLS allows cable operators to deploy L2 and L3 VPNs, as well as circuit emulation, across a packet-based network. In some cases, cable operators use MPLS to distribute multicast services although the data indicates that MPLS-based multicast distribution is not the norm today in cable.

PART II: ROUTING PROTOCOL USE CASES & ANALYSIS OF ROUTING TRENDS AMONG THE TOP 12 NORTH AMERICAN MSOs

Now that we have reviewed the most common routing protocols, let's discuss how these protocols are used in cable networks today. Starting with this core question: What is it, exactly, that makes an architecture specific to cable?

The answer is somewhat subjective, but perhaps the biggest single distinction is the CMTS (Cable Modem Termination System) access layer, and the IP-based services this access layer enables. Although the CMTS helps distinguish a cable service provider from other types of service providers, it is nonetheless intrinsically a router. Therefore it runs routing protocols.

Routing protocols and associated router-specific software features can be used to perform very complex operations, using advanced features. All North American MSOs are in general trying to accomplish the same goal, and therefore will utilize routing protocols in a similar way.

North of the CMTS, a cable operator's network typically resembles any service provider's network. However, some trends and themes do exist amongst cable operator networks. These themes or trends tend to develop out of general industry momentum, as well as close MSO engineering community interaction.

Multiple Systems Operators (MSOs) talk to each other at several engineering levels. Engineers and technical leaders tend to move within the cable community, bringing their methodologies with them. Peers from one cable company can often influence the routing protocol selection and design within another MSO. In cable, birds of a feather do flock together, creating trends that help shape the popularity and/or longevity of certain routing protocols.

The routing protocols required and deployed by an MSO will depend on their service offering. Services such as voice, data, video and commercial/business services are a common denominator among cable providers.

Voice and data services use the same IP network infrastructure and hence routing protocols. One distinction is that voice service is considered a high priority service requiring low latency and high availability. Considerations are taken within the routing protocol selection and design to guarantee those requirements. Both data and voice services are typically all unicast and therefore rely solely on unicast routing protocols. Routing protocols such as OSPF, ISIS, RIP and BGP are commonly used.

Video, on the other hand, is distributed in both the multicast and unicast formats, requiring the addition of a multicast routing protocol. These are the IGMP and PIM use cases.

Commercial business services often consist of the same services that exist on the residential side of the cable business. A main distinction in commercial services is the requirement of privacy. This requirement results in the need for routing protocols that can create virtual private networks for both layer 2 and layer 3 services. This, in turn, is the MPLS use case. The cable network routing design is a blending of all the above-mentioned protocols.

As networks continue to grow, it is not cost effective to build a separate network for each service, despite the fact that such separation does still exist in today's cable architectures. Today, the trend is to use one converged network for all services, including video distribution. Convergence continues to evolve across the industry. A cable company's size or geographic distribution plays a large role in the level of convergence deployed. This level of convergence also plays a role in how each routing protocols individual features might be utilized within the network e.g. multiple route processes.

Next we'll look at several of these use cases, to provide a base level of understanding about how each routing protocol is typically deployed.

OSPFv2, OSPFv3, & IS-IS v4/v6 IN CABLE ARCHITECTURES

Cable operators use IGPs in the same manner as other service providers: To distribute route reachability within their Autonomous System (AS.) What routes they advertise into their IGP, to distribute to the rest of the network, will likely depend on how they use BGP. BGP will be further explained later in this paper.

IGPs such as OSPF or ISIS are used in cable architectures to advertise the reachability of the network infrastructure. This includes all transit links and loopback addresses covering next hop address reachability for BGP. (Although RIP is not used within cable network architectures to distribute infrastructure reachability, it does have a cable use case and will be examined later.

Other commonly known IGPs, such as IGRP and EIGRP, have fallen out of favor compared to OSPF and ISIS, and have almost no usage within the cable community today. A few EIGRP networks still exist but they are typically a result of smaller unconverged and independent single service networks.

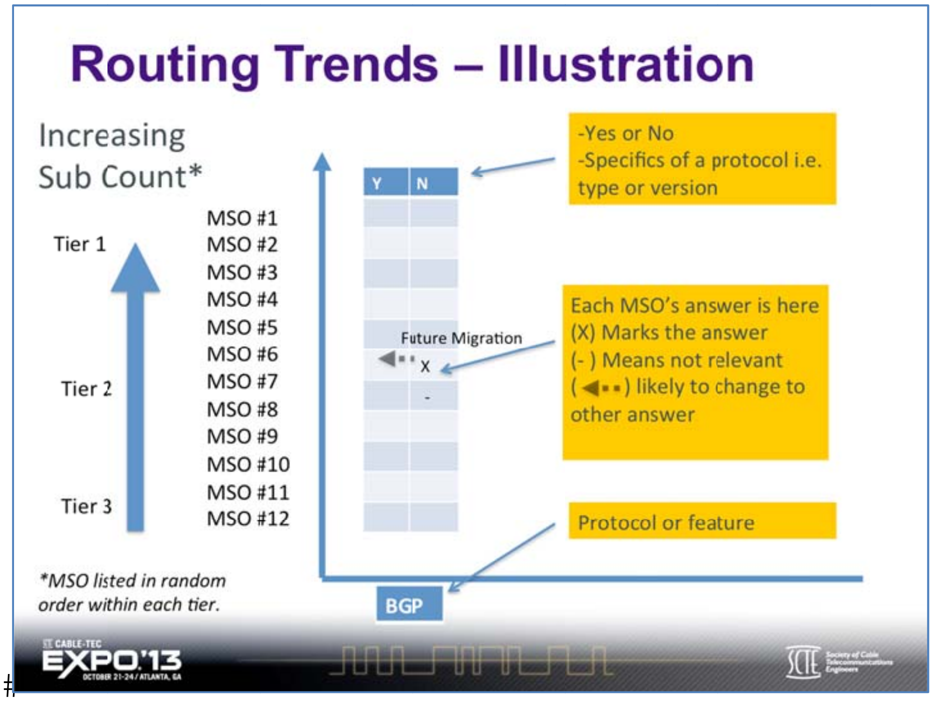
Customer route advertisement from a CMTS using an IGP will depend on how deep BGP is deployed within the network. In a scenario were BGP is not running on the CMTS, an IGP is used to advertise customer routes, typically to the next layer in the network. In this case, those networks are redistributed into BGP at the next routing layer. OSPF and/or ISIS route distribution within an AS is also used by multicast routing protocols such as PIM. Multicast routing protocols and how they utilize this IGP distributed information are discussed later.

In some cases, multiple OSPF processes are used to isolate service topologies. This use case requires some type of physical or logical topological separation. It is usually utilized when cable service providers want to segment certain types of traffic to certain links -- two links connected northbound from the CMTS, one for residential data, the other for commercial or specific to voice.

By default, routers perform destination-based routing, based on the routing information learned from a routing protocol. In some cases, cable operators don't want to use destination as a basis for choosing a next hop for a packet. In this case, a policy-based routing mechanism can be used to make a different forwarding decision. One example would be to make a forwarding decision based on the source address, which then forwards to a separate interface or logical tunnel. For these use cases, the goal is usually to separate services across different topologies.

As previously mentioned, and specifically for the purpose of this paper, we conducted an informal poll of the 12 North American MSOs, to ascertain what and how they utilize routing protocols. Figure 7 is intended to clarify how the MSO routing data within the survey was organized and displayed. The answer of a particular MSO is logged next to its name (anonymized) on the Y-axis. No specific order is utilized other than smaller MSOs are listed closer to the bottom of the graph.

The idea was to identify network protocol trends related to the size of an MSO. The general tier of an MSO is subjective and changes depending on the definition of tier used -- the standard service provider definition, or, size relative to other MSOs. It is used here to provide an approximation of overall subscriber size. The x axis or columns identify each question being collected. It might be a protocol, yes/no answer or architecture methodology.



Routing Trends – Illustration

-Yes or No
-Specifics of a protocol i.e. type or version

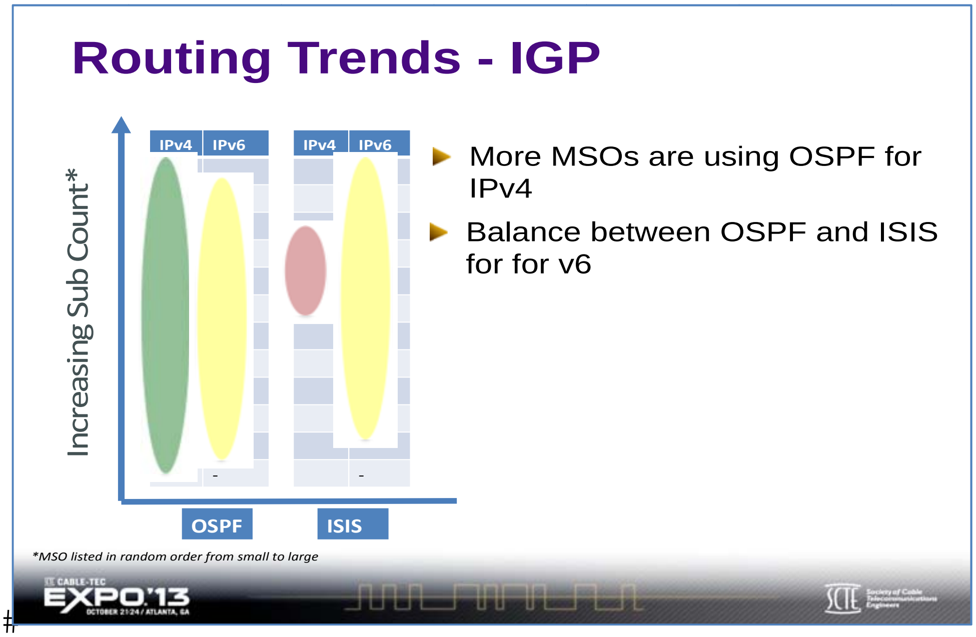
Each MSO's answer is here
(X) Marks the answer
(-) Means not relevant
(◀▪▪) likely to change to other answer

Protocol or feature

*MSO listed in random order within each tier.

Figure 7 An Overview of a Poll of 12 North American MSOs About Routing Protocols

The graph in Figure 8 illustrates that most MSOs (green highlight) are using OSPF for IPv4 route distribution vs. ISIS. Figure 8 also illustrates how OSPF and ISIS are distributed for IPv6. In this case there is a balance between all 12 MSOs.



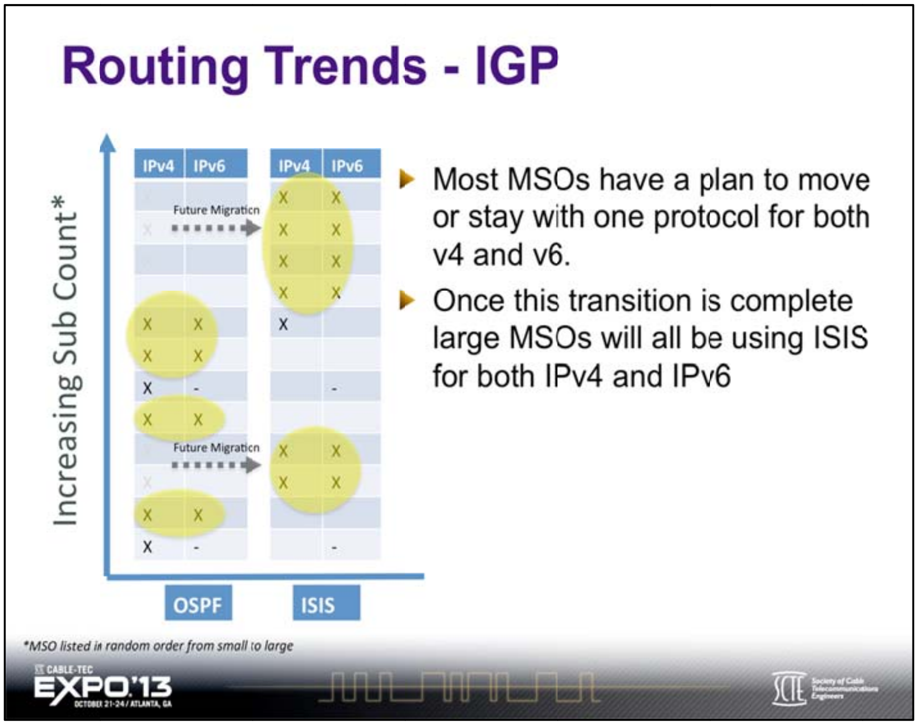
Routing Trends - IGP

- More MSOs are using OSPF for IPv4
- Balance between OSPF and ISIS for for v6

*MSO listed in random order from small to large

Figure 8 How 12 North American Cable Operators Are Using Interior Gateway Protocols

The same top 12 NA MSOs were also polled to determine their likely IGP direction and final migration plans for the next 12-18 months. For the MSOs running ISIS for IPv6, there was a desire and plan to migrate to MT-ISIS for both IPv4 and IPv6. When these MSO migration plans are taken into consideration (dotted arrows), the trending shows that one IGP protocol for both IPv4 and IPv6 will be utilized for most MSOs. This still maintains a balance across cable (yellow highlight) between the usage of OSPF and ISIS, but positions all large MSOs in the ISIS domain.



Routing Trends - IGP

- Most MSOs have a plan to move or stay with one protocol for both v4 and v6.
- Once this transition is complete large MSOs will all be using ISIS for both IPv4 and IPv6

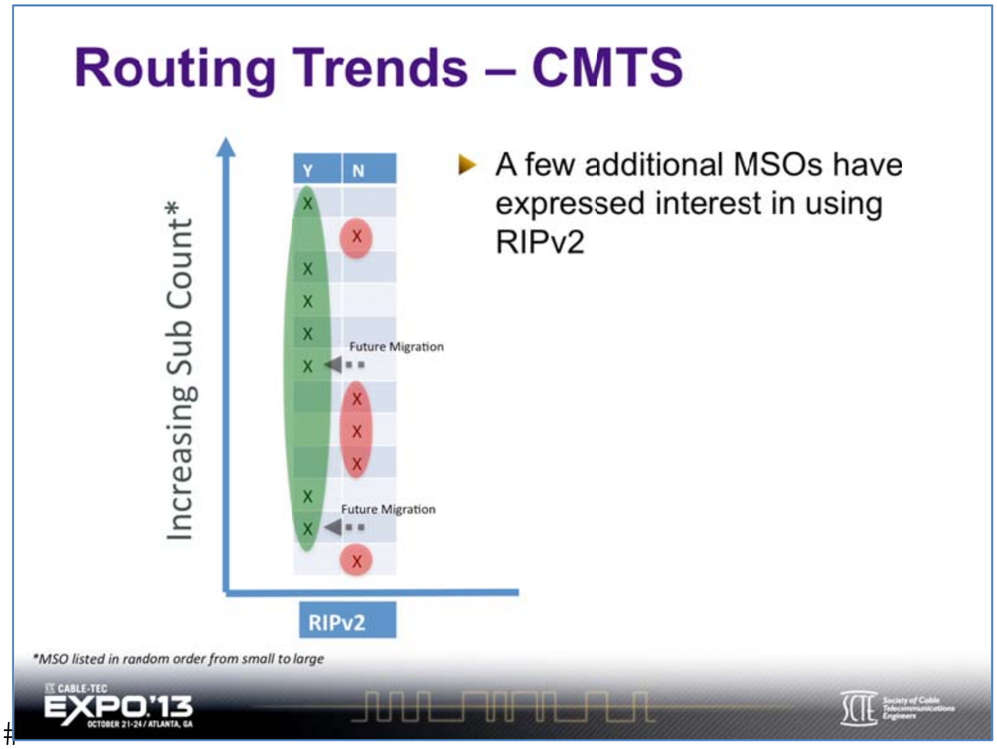
*MSO listed in random order from small to large

Figure 9 Most MSOs will use one IGP Protocol for both IPv4 and IPv6

GETTING TO KNOW RIP

RIPv2 is more widely used in cable then one might initially think. RIP is a simple and lightweight routing protocol. It is used today by some MSOs to dynamically learn subnet reachability information from customer routers connected to a CMTS. For this use case, a router or router with integrated cable modem runs RIP and advertises its subnets to the CMTS.

Once this information is learned, it is redistributed into either an IGP -- such as OSPF or ISIS, or BGP if used on the CMTS. RIP is only used as a small, one hop routing protocol and never across a large network. The use of static routing is typically the alternative for those MSOs who do not wish to use RIP.



Routing Trends – CMTS

- A few additional MSOs have expressed interest in using RIPv2

*MSO listed in random order from small to large

Figure 10 Routing Information Protocol (RIP) is Gaining Popularity

Once we factor in the MSOs' plans of record regarding the usage of RIPv2 we see that RIPv2 becomes the majority. MSOs using RIPv2 for IPv4 are also interested in using RIPng for IPv6 reachability information for the same use case.

COMMON MUTI-PROTOCOL BGP DESIGNS USED BY OPERATORS TODAY

BGP is used among service providers to essentially tell each other about the networks (subnets) they own. In this way, when a router within a service provider network receives a packet it knows whom it belongs to. Maybe it belongs to a local customer, maybe to another service provider across the country, or maybe to another network halfway around the planet.

BGP has become the Internet routing protocol standard for all service providers. Exactly how an MSO uses BGP varies depending on their size and requirements. At a high level, BGP is used to advertise a cable company's Internet-reachable routes to other service providers' routers.

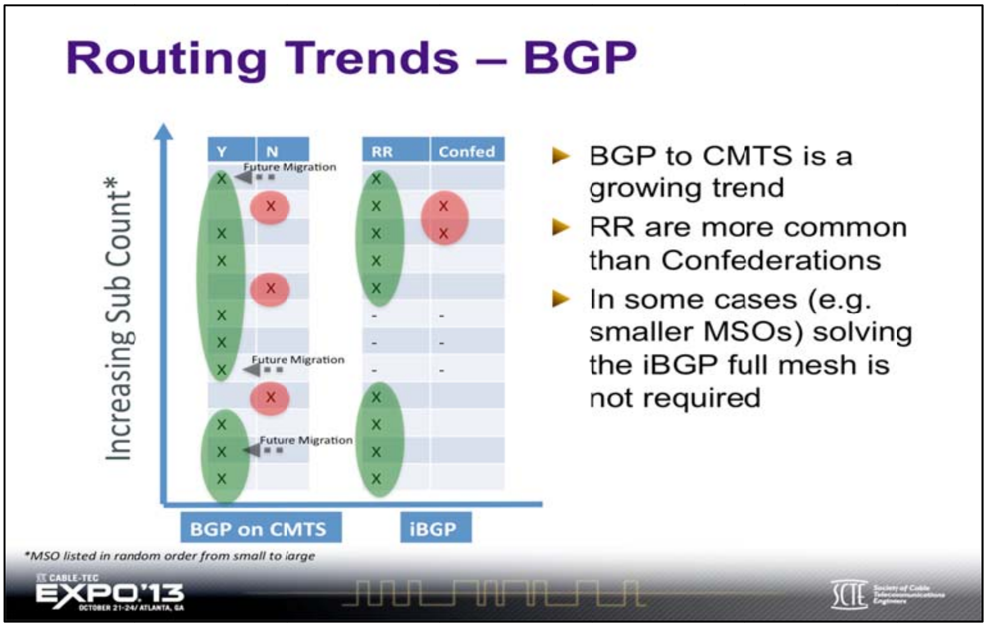
BGP embodies a rich set of policy features and enhancements, which give service providers control over what routes they learn, as well as what gets passed on to other companies. For example, attribute manipulation is a general and common practice among all cable operators to steer traffic in and out of their transit and peering interconnects.

One trend gaining momentum is the existence of BGP down to the CMTS. When BGP runs on the CMTS, subscriber routes and/or cable modem management addresses get distributed directly from the CMTS. Typically, only a BGP default route is sent to the CMTS, because a way out of the network is all that is required in most residential data scenarios. BGP communities are also used to tag advertisements for later manipulation and or identification.

When BGP is not used on the CMTS, an IGP is required to advertise customer routes to a higher layer within the network northbound of the CMTS. This requires redistribution and manipulation of the IGP routes, at this higher layer, in order to inject subscriber routes into BGP.

Multiprotocol BGP is commonly used in today's cable networks. A few large MSOs also use the capabilities of MPBGP to advertise IPv6, labels and multicast source address reachability.

Large BGP deployments often utilize a mechanism to minimize the full mesh requirement of iBGP. In some cases, usually amongst smaller MSOs or regional-only MSOs, only full mesh is used. Route reflectors (RRs) and confederations are both options. In Figure 11, one can see that the majority of MSOs use RR (green highlight) vs. confederations (red highlight). In some cases, two to three levels of route reflection are used.



Routing Trends – BGP

Y	N	RR	Confed
X		X	
	X	X	X
X		X	X
X		X	
	X	X	
X		-	-
X		-	-
X		-	-
	X	X	
X		X	
X		X	
X		X	

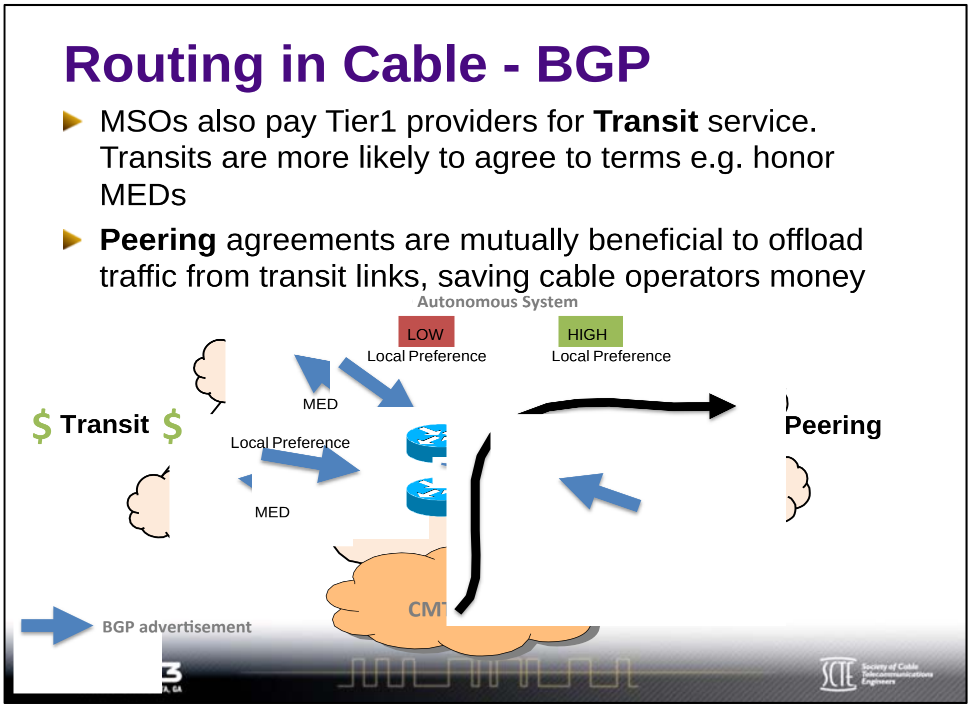
- BGP to CMTS is a growing trend
- RR are more common than Confederations
- In some cases (e.g. smaller MSOs) solving the iBGP full mesh is not required

*MSO listed in random order from small to large

Figure 11 Border Gateway Protocols Inside the CMTS

Figure 11 also shows the popularity of running BGP down to the CMTS access layer. The future migration of three MSOs (green highlight, left column) makes BGP on the

CMTS a very common future trend for almost all MSOs.



Routing in Cable - BGP

- MSOs also pay Tier1 providers for **Transit** service. Transits are more likely to agree to terms e.g. honor MEDs
- **Peering** agreements are mutually beneficial to offload traffic from transit links, saving cable operators money

Figure 12 The Business of BGP

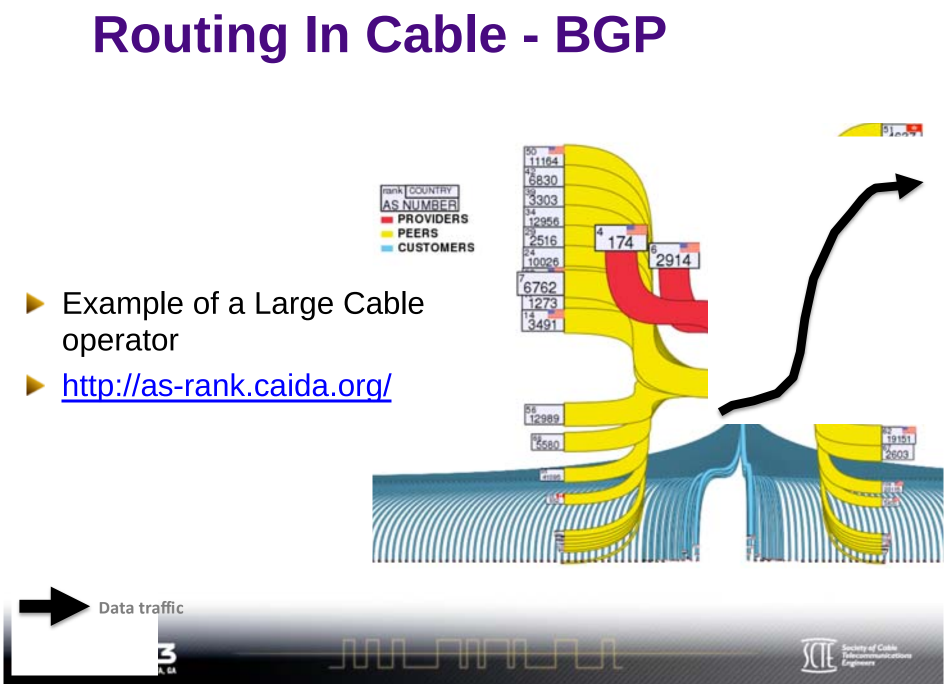
Like everyone else, service providers require network connectivity between their networks (it is the Internet, after all) in order to forward packets to their final destinations. These connections or data links run eBGP and can be defined as either transit or peering relationships. The distinction will change depending on what the link is used to do.

A transit link, for instance, is a pay-per-bit service connection, where all Internet routes are learned -- meaning that one can forward any packet across the link, and, for a price, the receiving service provider will forward that packet towards their final destination.

A peer link happens when two or more service providers mutually agree to exchange traffic through their autonomous systems. This is also called settlement-free. They will not act as a transit link and forward packets destined to networks, which they do not own.

All cable operators will have a transit connection (a paid tier provider) to provide Internet connectivity. Almost all large MSOs negotiate and build several peering agreements with different service providers. These agreements are mutually beneficial and not based on a cost-per-bit model, so a service provider is incentivized to use these peering links over any transit links.

This optimal traffic steering scenario provides a good business example of how cable operators utilize BGP attributes. BGP is used to lower the local preference attributes of BGP updates from transit providers, which causes traffic to prefer peering links (see Figure 12) saving MSOs transit expenses.



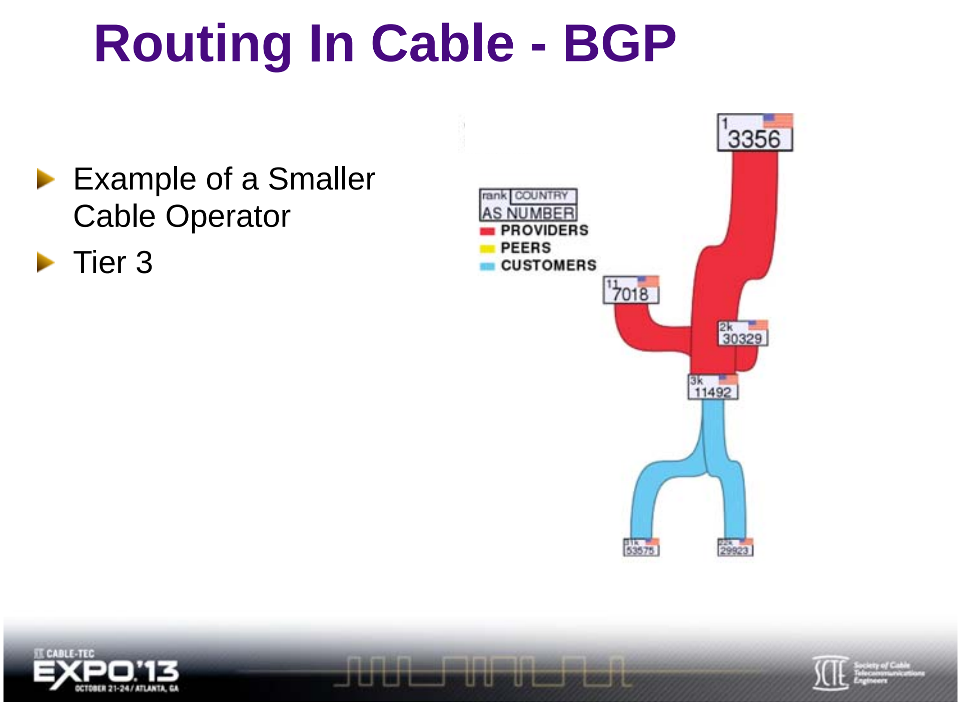
Routing In Cable - BGP

- Example of a Large Cable operator
- http://as-rank.caida.org/

Figure 13 How Large Operators Deploy BGP

The high level interconnectivity of the Internet, including how cable companies are interconnected, is public information. It can be helpful to look at your own company's BGP interconnectivity to better understand how traffic might flow out and into your network.

The Cooperative Association for Internet Data Analysis (CAIDA) is one way to do so. In its own words, CAIDA is "a collaborative undertaking among organizations in the commercial, government, and research sectors aimed at promoting greater cooperation in the engineering and maintenance of a robust, scalable global Internet infrastructure." Figure 13 shows an example of this public information, where transit service providers are shown as red links, peers as yellow and customers in blue.



Routing In Cable - BGP

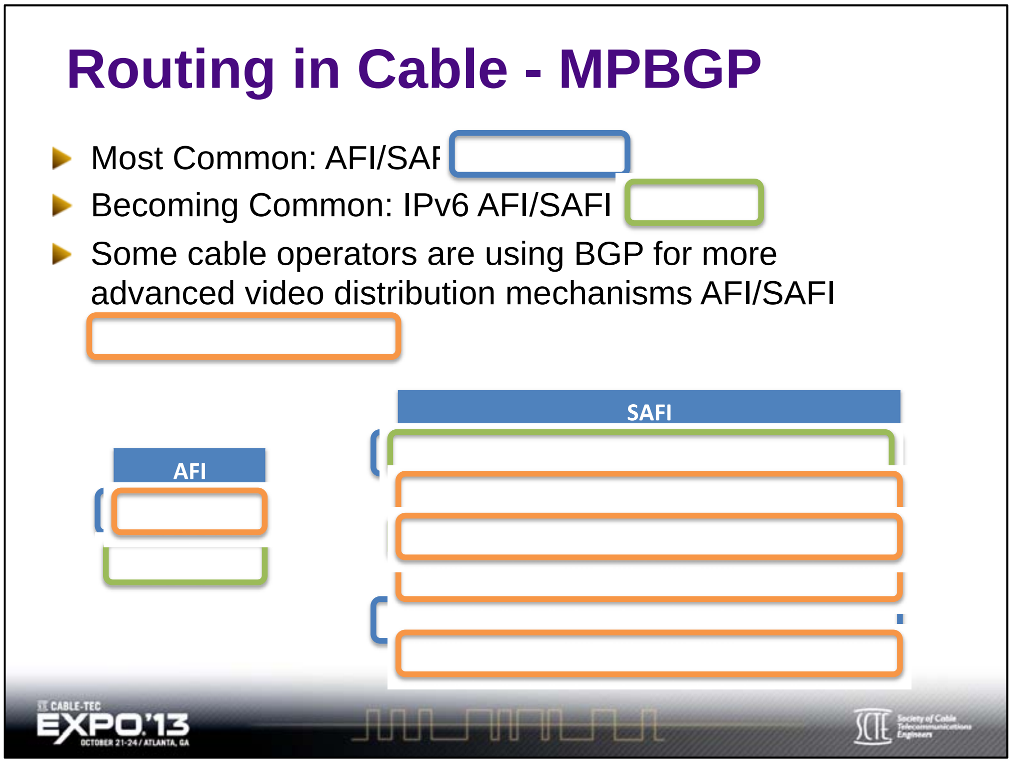
- Example of a Smaller Cable Operator
- Tier 3

Figure 14 How Small Operators Deploy BGP

Figure 14 is an example of a smaller MSO without established network peers. For this MSO all Internet traffic will use the red, provider-based AS connections.

Multi Protocol BGP has expanded the role of BGP and allows it to share many different types of information. This additional functionality is typically used inside an AS for the creation of complex services. MPBGP added the ability, for instance, for BGP to advertise reachability information for different address families. This capability hinges around the added BGP update fields of AFI (Address Family Identifiers) and SAFI (Subsequent Address Family Identifiers.)

<u>Commonly used AFI/SAFI values</u>
<u>AFI</u>
1 IPv4
2 IPv6.
<u>SAFI</u>
1 Unicast
2 Multicast - Source reachability
4 MPLS Label – NLRI with labels
5 mVPN -Distribute BGP Auto-Discovery and C-multicast Routes
128 MPLS-labeled VPN – VPN label
129 mVPN –Source reachability between PEs



Routing in Cable - MPBGP

- Most Common: AFI/SAF
- Becoming Common: IPv6 AFI/SAFI
- Some cable operators are using BGP for more advanced video distribution mechanisms AFI/SAFI

SAFI
AFI
CABLE-TEC EXPO'13 OCTOBER 21-24 / ATLANTA, GA
SCTE Society of Cable Telecommunications Engineers

Figure 15 Multi-Protocol BGP Deployments in Cable

MPBGP is common in cable architectures today, but is used differently depending on the size of the network and how services are transported. The most common in use today are 1/1 and 1/128 or basic IPv4 reachability and L3VPN service label distribution. The usage of dual stack (the running of both IPv4 and IPv6) has introduced the MPBGP AFI/SAFI 2/1 and 2/4. 2/4 is used to support 6PE for those operators running an MPLS backbone. BGP is also being used for more advanced multicast VPNs. 1/2 , 1/5 , and 1/129 are all used to support a multicast-based VPN. AFI/SAFI value of 1/2 is also used in native multicast forwarding to protect multicast source addresses by removing them from the main unicast Internet routing table.

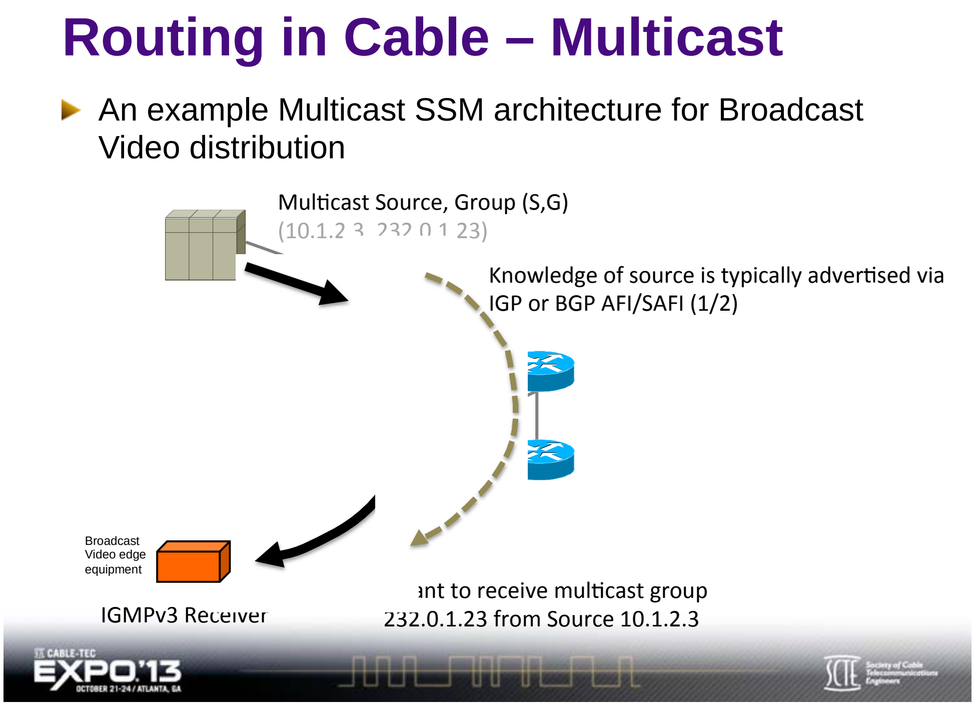
HOW MULTICAST ROUTING PROTOCOLS ARE USED IN CABLE ARCHITECTURES

The original cable service -- television -- has evolved enormously through the decades. It has become both a real-time and time-shifted on demand service, requiring different routing architectures and protocols. Multicast routing is a requirement today for most video distribution solutions within cable. And if it's not, it will be.

Multicast routing is performed with a combination of an Interior Gateway Protocol (IGP), Protocol Independent Multicast (PIM), and Internet Group Management Protocol (IGMP.) PIM and IGMP are the workhorses, while IGP plays more of a supporting role. There are other protocols that can be used to create multicast routing state in routers, but PIM with IGMP are the two most common in cable.

Several variants of PIM exist, each with a specific use case. One example is PIM-ASM, where the "ASM" stands for "Any Source Multicast." This protocol is typically paired with IGMPv2 and has fallen out of favor for most cable deployments. Rather, PIM-SSM (Source Specific Multicast) is typically paired with IGMPv3 and is the common IP-packet based multicast routing protocol used within cable today.

The most common use case of multicast routing in cable is for video distribution.   Many MSOs have or are moving from legacy distribution of video to an IP, packet-based distribution of video. In some cases MPLS based protocols are used to transport multicast traffic, although in cable today, this is not as common as packet-based distributions. Although not an exhaustive list, the other common use cases for multicast routing in cable are DSG (DOCSIS Signaling Gateway) traffic and VOD library content distribution.


Routing in Cable – Multicast

- An example Multicast SSM architecture for Broadcast Video distribution

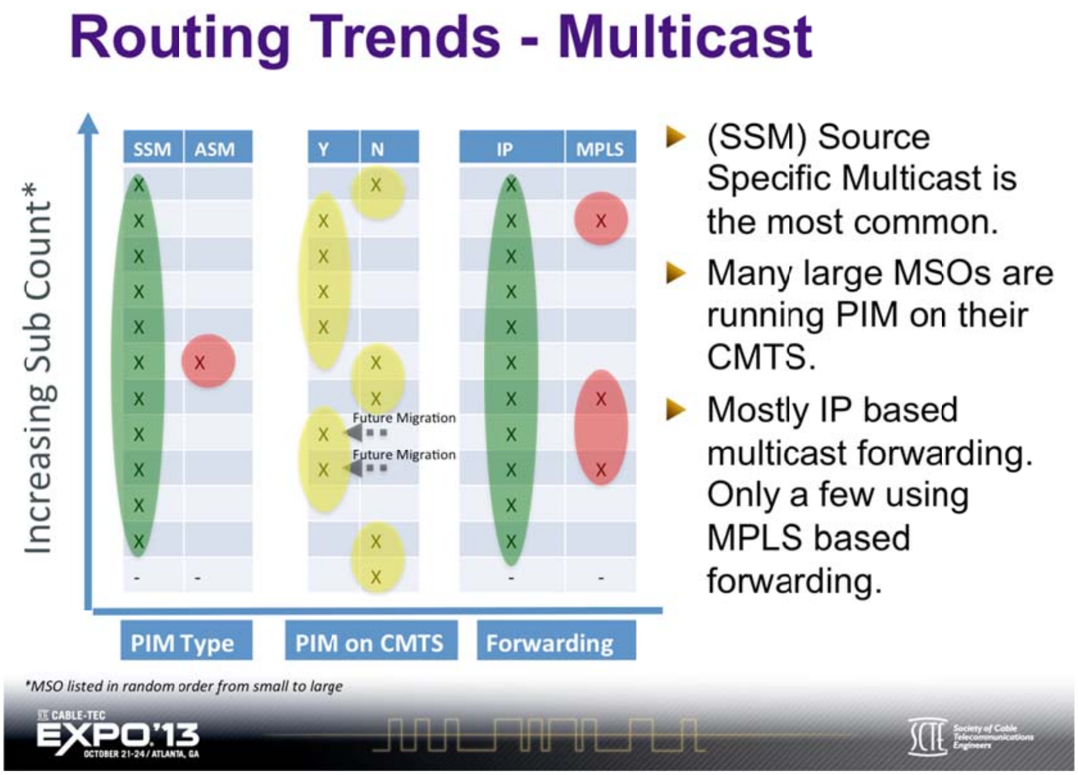
Multicast Source, Group (S,G)
(10.1.2 3, 232 0 1 23)
Knowledge of source is typically advertised via IGP or BGP AFI/SAFI (1/2)
Broadcast Video edge equipment
IGMPv3 Receiver
ınt to receive multicast group 232.0.1.23 from Source 10.1.2.3
CABLE-TEC EXPO'13 OCTOBER 21-24 / ATLANTA, GA
SCTE Society of Cable Telecommunications Engineers

Figure 16 The Source-Specific Multicast

Figure 16 demonstrates how an Interior or Border Gateway Protocol, plus IGMP and PIM, work together to build multicast forwarding state in a router. Shown is an example of PIM-SSM with IGMPv3. The first hop router that receives the IGMP request uses PIM to signal the source (orange arrows). This first hop router knows where to find the source, either using an IGP like OSPF or ISIS, or it uses BGP via the multicast source AFI/SAFI (dotted brown arrow).

Once a multicast path is established and every router along the path knows how to handle this particular source and group address combination (10.1.2.3,232.0.1.23), the multicast (black arrow) will flow from source to receiver.

In Figure 17, one can see that there emerges a clear winner when it comes to PIM-type routing used in cable today. Only one case of PIM-ASM was found and was attributable to legacy equipment. The preference for the majority of North American MSOs is PIM-SSM with IGMPv3, which provides a much simpler and secure multicast routing solution.



Routing Trends - Multicast

- (SSM) Source Specific Multicast is the most common.
- Many large MSOs are running PIM on their CMTS.
- Mostly IP based multicast forwarding. Only a few using MPLS based forwarding.

*MSO listed in random order from small to large

Figure 17 Source-Specific Multicast Is Widely Used in Cable

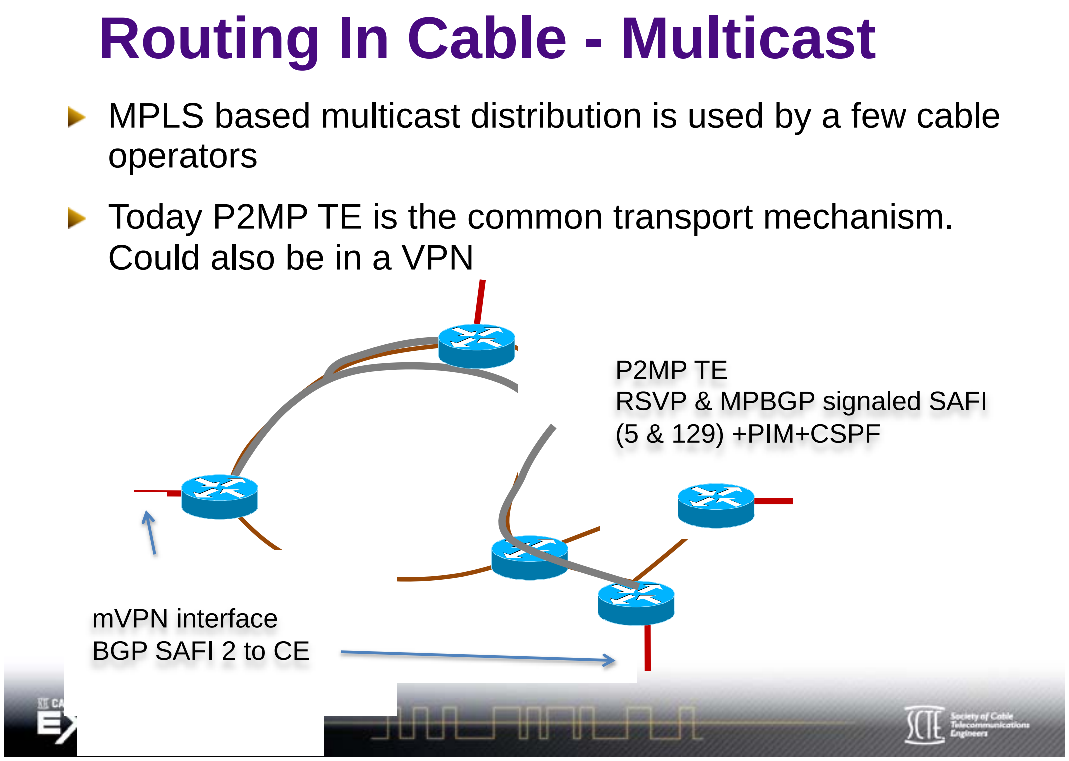
The popularity of PIM-SSM has also created a use case for static IGMP joins. This is because a lot of legacy equipment doesn't yet support IGMPv3. A static join is used on the router and works by statically emulating what an IGMP protocol would dynamically perform. The limitation is that the static join is always present -- however, it's rarely an issue since the end device usually always wants to receive the multicast content.

Figure 17 also shows that almost all 'large' cable operators are running PIM on their CMTS. An additional few are in tests, with plans to migrate to PIM on the CMTS. In general however, across all 12 North American MSOs, there will still be a balance (yellow highlight) among those who do and do not run PIM on the CMTS.

Although today the possibility to distribute multicast traffic via an MPLS or label-based environment exists, most cable operators still choose to use packet-based forwarding mechanisms, as shown in the last column of Figure 17. For those that do use an MPLS mechanism they also still use a packet-based scheme.

Point-to-Multipoint Traffic Engineering (P2MP TE) is typically used as the transport for those few operators using an MPLS-based forwarding network to forward multicast packets. PIM, RSVP and MPBGP are used as control protocols.

The below illustration details how an MPLS network can be used to distribute Multicast traffic. Several MPBGP address families (AFI/SAFI values) are used to distribute the necessary information.



Routing In Cable - Multicast

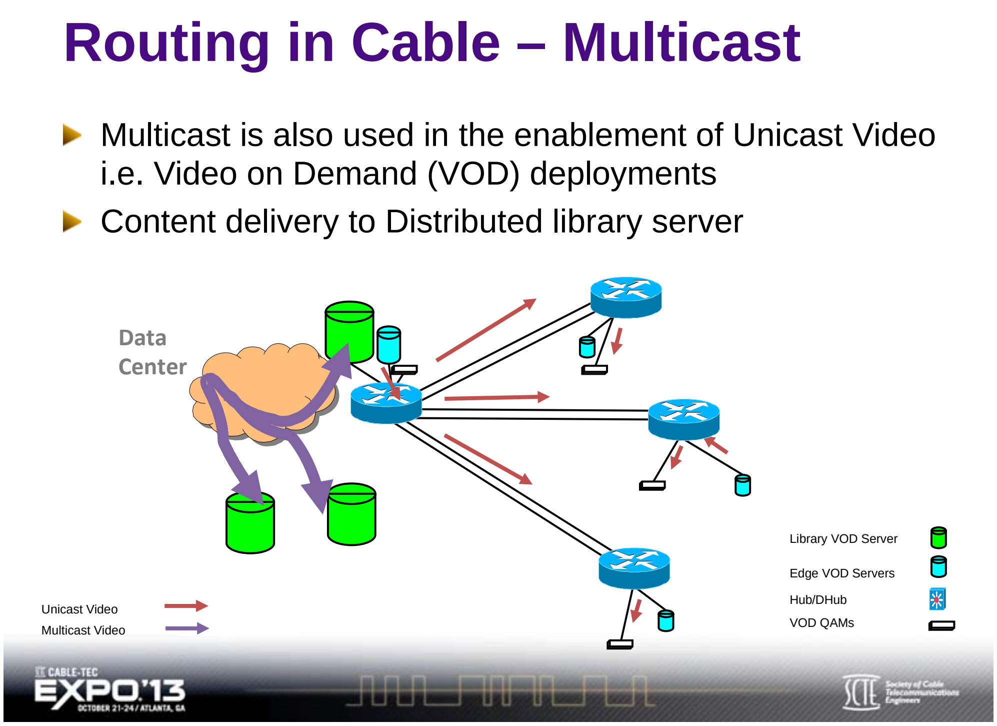
- MPLS based multicast distribution is used by a few cable operators
- Today P2MP TE is the common transport mechanism. Could also be in a VPN

P2MP TE
RSVP & MPBGP signaled SAFI
(5 & 129) +PIM+CSPF

mVPN interface
BGP SAFI 2 to CE

Figure 18 MSOs Using MPLS-based Multicast Use Point-to-Point Traffic engineering (P2MP-TE)

Another use case for the multicast protocol within a cable network is to support unicast video deployments or VOD solutions. Traditional, QAM-based VOD traffic is unicast. However, that content must get distributed to multiple distributed library servers across a network. Multicast is used to distribute this content (purple arrow) Figure 19.



Routing in Cable – Multicast

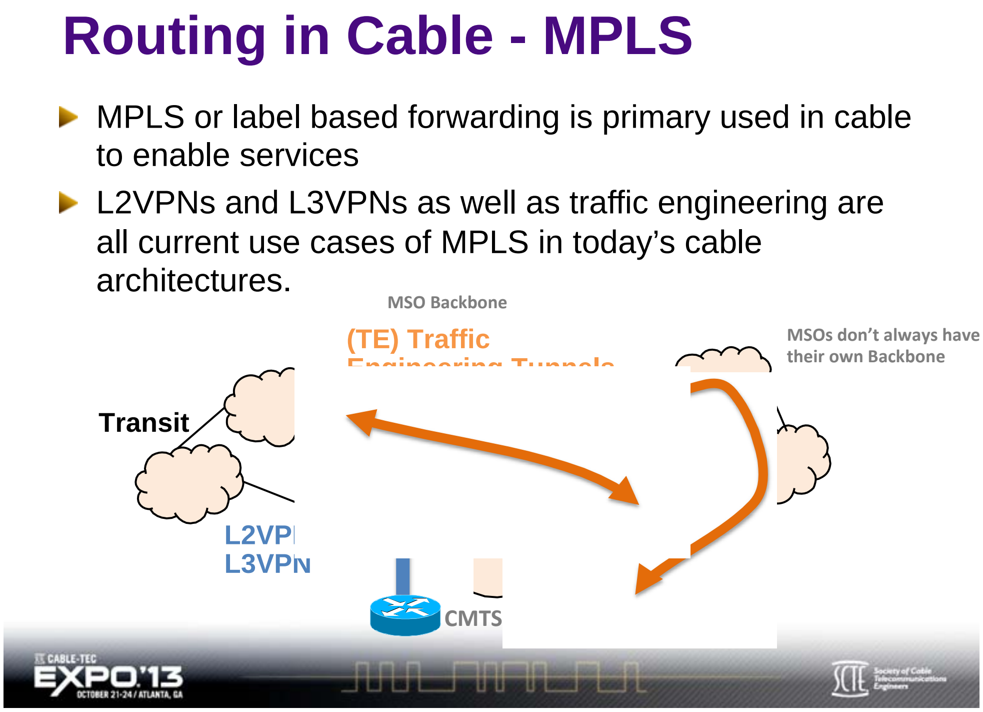
- Multicast is also used in the enablement of Unicast Video i.e. Video on Demand (VOD) deployments
- Content delivery to Distributed library server

Data Center
Library VOD Server
Edge VOD Servers
Hub/DHub
VOD QAMs
Unicast Video
Multicast Video
CABLE-TEC EXPO'13 OCTOBER 21-24 / ATLANTA, GA
SCTE Society of Cable Telecommunications Engineers

Figure 19 The Importance of Multicast for IP-based VOD

THE ROLE OF MPLS IN CABLE

MPLS has become the go-to protocol for carving out services across a network infrastructure. It is used as a forwarding service in several cable networks. Specifically, MPLS is used to enable Layer 2 and Layer 3 VPN services as well as to optimize traffic patterns through traffic engineering. Traffic Engineering through MPLS avoids destination-based, connection-less routing and enables a high level of control over how traffic is transported across a network. A typical use case for smaller MSOs (Figure 20) is the ability to steer traffic across leased circuits. For larger MSOs, the ability to optimize infrastructure and/or minimize transport redundancy is a primary benefit.

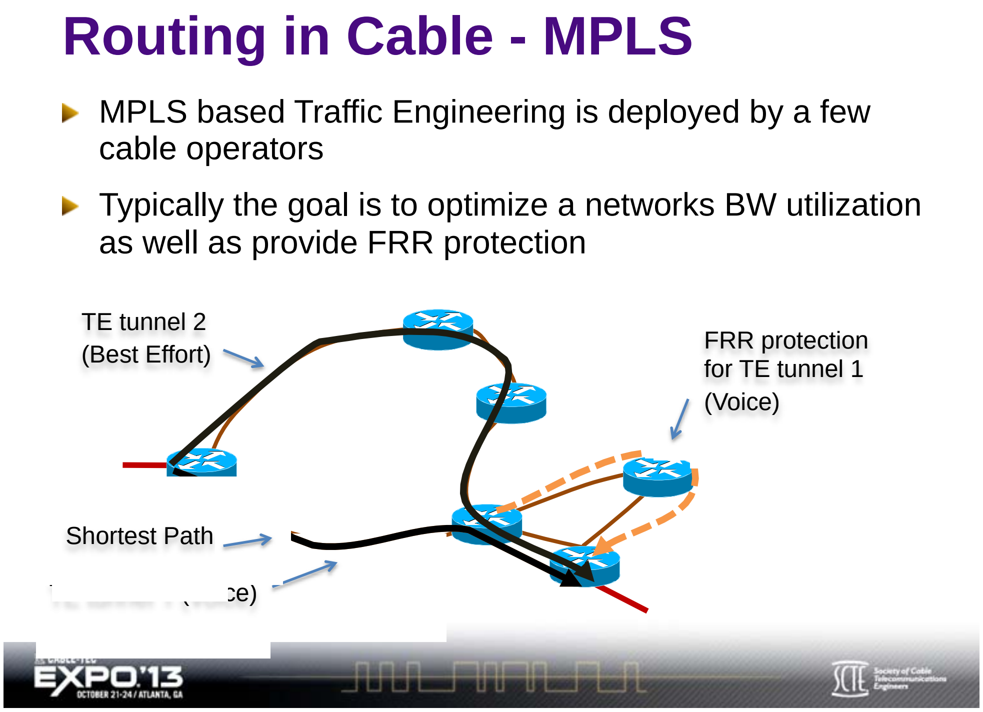


Routing in Cable - MPLS

- MPLS or label based forwarding is primary used in cable to enable services
- L2VPNs and L3VPNs as well as traffic engineering are all current use cases of MPLS in today's cable architectures.

Figure 20 Use Cases for Cable MPLS Services

Figure 21 is a high-level use case of how Traffic Engineering in MPLS is used. Fast Reroute is typically used to provide fast convergence for voice services.

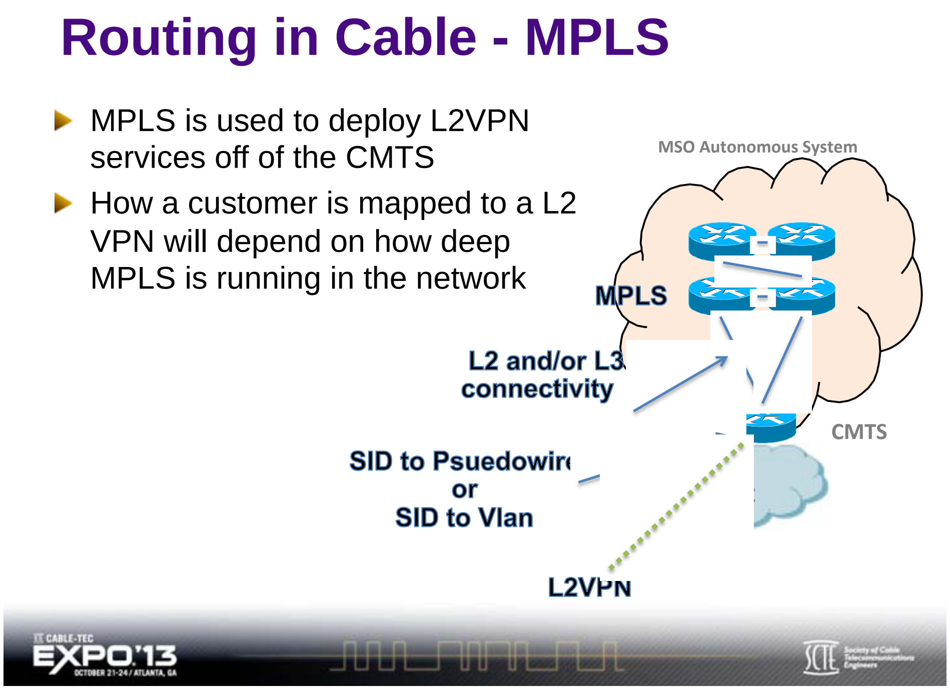


Routing in Cable - MPLS

- MPLS based Traffic Engineering is deployed by a few cable operators
- Typically the goal is to optimize a networks BW utilization as well as provide FRR protection

Figure 21 Traffic Engineering Gains Make MPLS a Way to Optimize Bandwidth

Some cable operators use MPLS to enable L2VPN services from the CMTS shown in Figure 22. MPLS is not required for this service, however, it's often used to some extent even if it's only running north of the CMTS. If MPLS is running on the CMTS, then a Service Identifier to Psuedo Wire mapping takes place (SID to PW mapping). If no MPLS is running on the CMTS, then a Service Identifier to Virtual Local Area Network mapping (SID to VLAN mapping) takes place. In this case, L2 connectivity is used to transport northbound VLANs, where they are then either terminated on an MPLS-based PE, or trunked further to their destination.



Routing in Cable - MPLS

- MPLS is used to deploy L2VPN services off of the CMTS
- How a customer is mapped to a L2 VPN will depend on how deep MPLS is running in the network

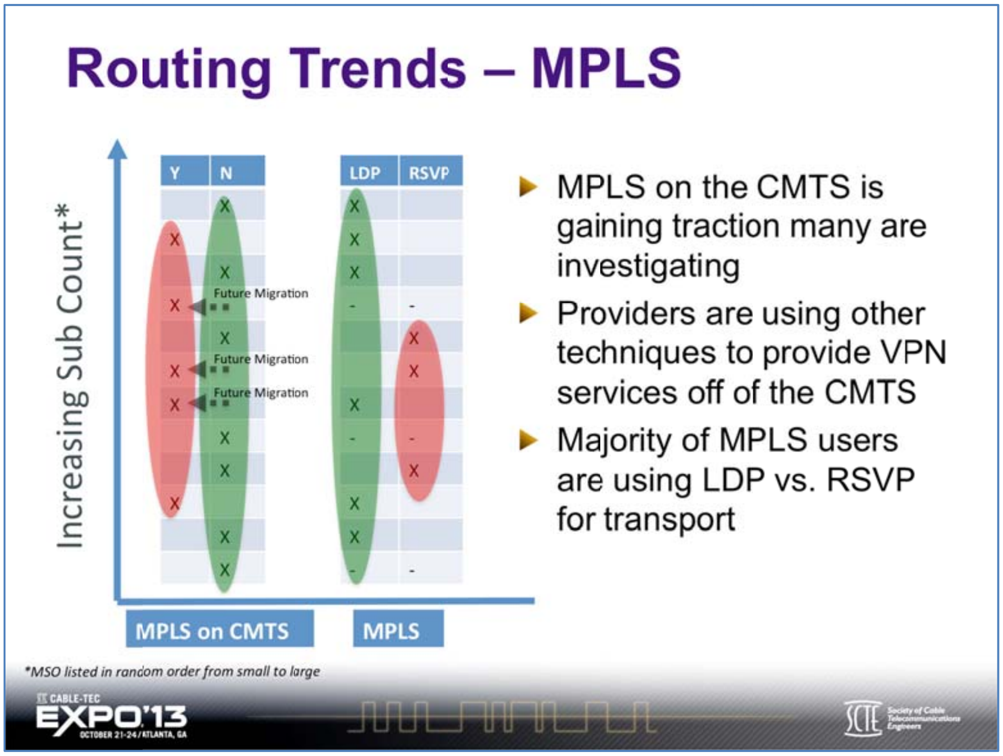
MSO Autonomous System
MPLS
L2 and/or L3 connectivity
CMTS
SID to Psuedowir
or
SID to Vlan
L2VPN
CABLE-TEC EXPO'13 OCTOBER 21-24 / ATLANTA, GA
SCTE Society of Cable Telecommunications Engineers

Figure 22 MPLS Routing for Layer 2 VPN Services

The main purpose of Figure 23 is to highlight which MSOs are utilizing some level of MPLS forwarding. The illustration also identifies those MSOs using traffic engineering, such as RSVP signaling.

Today the majority of MSOs that have decided to utilize MPLS forwarding use LDP based label distribution (green highlight) Traffic Engineering using RSVP is only used by a few cable operators (red highlight.)

MPLS on the CMTS is being investigated by many MSOs, but today it is not the majority (red highlight).



Routing Trends – MPLS

Y	N	LDP	RSVP
	X	X	
X		X	
	X	X	
X	Future Migration	-	-
	X		X
X	Future Migration		X
X	Future Migration	X	
	X	-	-
	X		X
X		X	
	X	X	
	X	-	-

- MPLS on the CMTS is gaining traction many are investigating
- Providers are using other techniques to provide VPN services off of the CMTS
- Majority of MPLS users are using LDP vs. RSVP for transport

*MSO listed in random order from small to large

Figure 23 MPLS Routing use case

SUMMARY

When it comes to network protocols, today's cable operators have to deal with everything. Multiple routing protocols are tested, deployed and managed. As an industry, cable providers do tend to work towards common solutions, but network size is typically a large influence in the final architectural and protocol direction they choose.

The data we gathered from the top 12 North American MSOs shows that the trend for the majority is to move towards the same IGP for both IPv4 and IPv6, although a balance between the use of OSPF and ISIS still exists.

RIPv2 is gaining momentum and will soon be used at the CMTS access layer, in the majority of MSO networks.

Route reflector solutions are the majority within cable with only a few confederation solutions deployed.

Packet-based forwarding for multicast is still the majority-forwarding scheme for cable networks. Most label-based forwarding activities are using LDP vs. RSVP as a transport mechanism.

Finally, routing protocols such as BGP, PIM and MPLS deployed on the CMTS are gaining momentum.

Cable networks are constantly evolving to accommodate new services and business models. Understanding how routing protocols are used provides a knowledge base from which to work as these protocols evolve and their use cases morph to enable further new cable-delivered services.

Abbreviations and Acronyms

ABR: Area Border Router – A router type within the Open Shortest Path First (OSPF) protocol defined by its position in the network and what areas it touches.

ASBR: Autonomous System Boundary Router – A router type within the Open Shortest Path First protocol defined by its position in the network.

BGP: Border Gateway Protocol – A path vector protocol that maintains path information to avoid loops. The only accepted exterior gateway protocol (EGP).

CAIDA: The Cooperative Association for Internet Data Analysis. A collaborative undertaking among organizations in the commercial, government, and research sectors aimed at promoting greater cooperation in the engineering and maintenance of a robust, scalable global Internet infrastructure.

CMTS: Cable Modem Termination System. The “headend part” of communications between cable modems and the Internet.

iBGP: Internal Border Gateway Protocol - An identifier of the type of peer establishment between routers, where the “i” stands for “internal.” Two peers within the same autonomous system are considered iBGP.

eBGP: External Border Gateway Protocol – It identifies the type of peer establishment between routers. Two BGP peers from differing autonomous systems are considered eBGP.

IGMP: Internet Group Management Protocol – A method for clients to advertise their interest in a particular multicast group. IGMP is a client protocol used to dynamically signal the first hop router of the interest in receiving a particular multicast group.

IGMPv3: Internet Group Management Protocol, version 3 – A dynamic method for clients to advertise their interest in a particular multicast group, by including the source as well as the group address.

IGMPv2: Internet Group Management Protocol, version 2 – A method for clients to advertise their interest in a particular multicast group, by group address, not source. This forced the routing network to determine who the source was, which significantly increased the complexity of multicast routing complexity.

IGP: Interior Gateway Protocol – a method of distributing routing information within an autonomous system or single operational entity. Designed to be fast and efficient, with minimal policy enforcement. Examples used in cable include but are not limited to OSPF, ISIS and RIP v2.

ISIS: Intermediate System to Intermediate System – a link state routing protocol that uses a 2-level hierarchy. ISIS routers are designated as level 1 (intra-area), level 2 (inter-area) or level 1-2 (both).

LDP: Label Distribution Protocol: A protocol used for distributing labels in an MPLS environment. LDP relies on the underlying routing information provided by an Interior Gateway Protocol in order to forward label packets. These labels are associated with routes, attached to packets and used my MPLS for forward those packets in an MPLS environment.

LSA: Link State Advertisement: An LSA is a type of packet sent by a router running OSPF that contains and shares a routers database of link state information. It is the mechanism a router uses to distribute information about a router's link state to its OSPF peers.

MP-BGP: Multi-Protocol Border Gateway Protocol: An extension of a BGP's capabilities beyond the advertising of unicast, IPv4-based reachability. Companion terms include "AFI" (Address Family Identifier) and "SAFI" (Subsequent Address Family Identifier," which in combination identify the set of network layer protocols to which the address carried in the next hop field must belong.

MLD: Multicast Listener Discovery – A method used by IPv6 routers and similar in concept to IGMP for IPv4. In MLD, the protocol is embedded in ICMPv6 instead of using a separate protocol. MLDv1 is similar to IGMPv2 and MLDv2 similar to IGMPv3.

MPLS: Multiprotocol Label Switching -- A routing protocol that assigns labels to routing information and distributes those labels. Those labels become the mechanism to make forwarding decisions, and are typically derived from the knowledge of other routing protocols.

Multicast Routing Protocols: A method for creating distribution trees that are typically (but not always) based on a demand model. If you seek this-or-that multicast service, you ask for it (create a demand) and the multicast routing protocol will build a forwarding path to you.

NLRI: Network Layer Reachability Information – a term used in BGP discussions to characterize the information within BGP messages that describes reachability.

OSPF: Open Shortest Path First – A link state routing protocol that distributes subnet reachability to everyone within an area, until all participating nodes have identical databases (network knowledge). The Shortest Path First algorithm is then run independently on each router. OSPF uses Internet Protocol for router updates, like Link State Advertisements, or LSAs.

P2MP TE: Point to Multi-Point Traffic Engineering – A one to many type of traffic engineering tunnel. Used to transport multicast packets.

Peer link: A linkage of two or more service providers, which have mutually agreed to exchange traffic through their autonomous systems.

PIM: Protocol Independent Multicast – a method of building “trees” of reachability between multicast listeners and multicast sources. PIM works by using any underlying Internal Gateway Protocol to build reachability trees.

PIM-ASM: Protocol Independent Multicast – Any Source Multicast. A protocol typically pared with IGMPv2 that is seldom used in cable deployments.

PIM-SSM: Protocol Independent Multicast - Source Specific Multicast. A method of forwarding multicast state information that is built within a router, based on both the source and destination addresses. Expressed as (S,G) and pronounced “S comma G,” where S stands for source IP address and G stands for group IP address. PIM-SSM is typically pared with IGMPv3 and is the common IP packet-based multicast routing protocol used within cable today.

RIP: Routing Information Protocol – A simple, lightweight, distance vector protocol that works by sending its neighbor only the reachability information it knows of. Hop count limited to 15 hops from source to destination.

RIP v2: Routing information Protocol version 2 was enhanced to included the ability to carry subnet information

RR: Routing Reflector – a method used for BGP deployments to avoid the need for a full-mesh deployment. One BGP speaking router is defined as the Route reflector and all other BGP routers only peer to it vs. each other.

SID to PW mapping: Service Identifier (SID) is used in DOCSIS to identify a CPE. The SID is mapped to an MPLS based Pseudo Wire on the CMTS so that a L2VPN service can be established to that CPE.

SID to VLAN mapping: Service Identifier (SID) is used in DOCSIS to identify a CPE. The SID is mapped to a VLAN so that a broadcast domain or L2VPN can be established to that CPE.

Transit Link: A pay-per-bit service connection, where all Internet routes are learned -- meaning that one can forward any packet across the link, and, for a price, the receiving service provider will forward that packet towards their final destination.