Making multicast work in IP networks

A credible concept for satisfying IPTV and beyond, proofed by the ResIP Center

Nokia Siemens Networks

White Paper

ResIP certified

Executive summary

Nokia Siemens Networks develops and implements carrier-grade, end-to-end IP-powered (Internet Protocol) solutions featuring best-in-class quality of service (QoS), reliability, management, and security. These IP solutions for voice, video, and data applications first undergo a battery of rigorous proof-of-concept tests before they earn this 'best-in-class' rating. To this end, the company has put in place one of the world's most advanced proving grounds, the Resilient IP Center, or ResIP Center for short.

The ResIP Center takes a holistic approach, factoring the entire IP network into the equation when validating solutions. Combining meticulous attention to detail with the ability to step back and see the larger picture, the ResIP Center optimizes and certifies solutions to minimize technical deployment risks and ensure carrier-grade interoperability, resiliency, QoS, and scalability. What's more, the Center's solutioneers incorporate not only best-of-breed products from Nokia Siemens Networks, they also cherry-pick the finest products of leading IP/Ethernet partners to accelerate time to market and guarantee an assured user experience.

The company also provides professional network planning and implementation services to help service providers (SPs) tailor ResIP-certified solutions to suit their networks and customize their multicast capabilities and IPTV solutions.

This paper looks closer at what it takes to make multicast work, including:

- How networks must be adapted to satisfy the demanding bandwidth, reliability, responsiveness, and QoS requirements of IPTV services
- What types of access network architectures will get the job done
- Layer 2 control and how it can improve dynamic QoS mechanisms
- How Protocol Independent Multicast – Sparse Mode (PIM-SM), Protocol Independent Multicast – Source Specific Multicast (PIM-SSM), and Point-to-Multipoint Label Switched Paths (MPLS LSPs) can serve to distribute IPTV content through the IP backbone

The case for multicast

Growing competition and fixed-to-mobile substitution have taken their toll on the voice service revenues of telecom operators around the world. To make matters worse, cable TV operators continue to encroach on carriers' territory by offering competing voice services.

On the upside, studies show that customers are willing to pay for an attractive package of services comprising voice, video, and data. Most prefer to deal with a single trusted SP who furnishes and bills for all the services they need. All this is compelling incumbent wireline carriers to rethink their strategy and realign their business. The time to transition from a yesteryear's stance as a mere pipeline provider to become a one-stop, full-service provider is now.

To make this new business model work, carriers have to satisfy users' expectations for multimedia services. Their networks must be geared up to deliver voice, data, video-on-demand (VoD), and TV broadcast services, smoothly and reliably. In the final reckoning, the success of such a new business model hinges on the quality of the user experience.

This business opportunity is enticing, but the technical barriers are high and the risks can be considerable. As some misfortunate SPs can attest, unresolved technical issues can stall projects, force them to roll out incomplete offerings, and diminish QoS in unexpected ways. Customers who experience such poor service take a dim view of triple-play offerings and tend to look elsewhere for better service. It does not take an accountant's acumen to gauge the impact on SPs' margins.

The good news is that with the right engineering skills, these barriers can be negotiated and risks mitigated. Nokia Siemens Networks has joined forces with the routing experts Cisco Systems and Juniper Networks to develop end-to-end solutions geared to help SPs roll out and capitalize on VoIP, triple-play, and IPTV services. With all this expertise pooled in one place, the ResIP Center is able provide to SPs validated, optimized, and certified IP solutions that deliver an assured quality of experience.

Users' great expectations for IPTV

Rolling out new services is always about managing business and technical risks; achieving success is about meeting consumer needs and expectations. Having enjoyed terrestrial, cable, and satellite TV services for decades, people have very high expectations indeed when it comes to availability, latency, and quality. If IPTV services are to survive and thrive, IP networks must be engineered to deliver guaranteed QoS even under adverse network conditions.

Users will not settle for anything less than what they already take for granted. This compels SPs to:

- Deliver assured quality of service in real time, with fast interactive response and minimum video and audio glitches
- Offer a wide range of video services such as broadcast TV, VoD, personal video recording, time-shifted viewing, and premium content
- Warrant highest availability; that is, rule out any single point of failure and put in place fast recovery mechanisms for sub-second failover times
- Ensure easy and convenient handling

The need to meet users' expectations is one side of the business equation. The other is the necessity to contain the operational and capital expenditure (OPEX and CAPEX) that squeezes margins. To this end, SPs must find ways to:

- Make the most of network resources for both unicast and multicast applications
- Deploy advanced management and zero-touch provisioning tools to minimize operational effort
- Bring new services to market that much quicker
- Remain open and flexible for the many different access types, network architectures, aggregation technologies, and protocols while protecting investments in legacy assets

Gearing networks up for IPTV

IPTV and VoD services pose some daunting challenges for today's broadband IP networks. This section looks closer at what these challenges are and how they can be mastered with the applied design guidelines and engineering efforts of the ResIP Center.

The bandwidth dilemma

While IPTV services have a healthy appetite for bandwidth, VoD services' craving for network resources is voracious. What's more, their time-sensitive nature mandates highly predictable performance. Depending on compression and coding technology, the network must be able to deliver the following transmission rates:

- 3.5-5 Mbit/s for MPEG-2-coded SD VoD streams or IPTV stream per TV channel
- Up to 2 Mbit/s for H.264-coded (MPEG-4 part 10) SD VoD streams or IPTV stream per TV channel:
- 8 to12 Mbit/s for HD signals coded with H.264

IPTV and video services are not created equal when it comes to consuming network resources, and bandwidth demand is just one of several factors that have to be considered before implementing an IPTV service.

What IPTV services need from the network

The overall bandwidth requirement depends on the total number of streamed IPTV channels. In other words, the aggregate transmission rate of IPTV content measured in Mbit/s equals the sum of all concurrent streams. Each IPTV channel is sent just once from the video head-end to the network, regardless of how many addresses may be receiving the TV broadcast. Multicast implementations in the core and access networks distribute the signals to all subscribers. Say an SP is broadcasting 80 IPTV channels, each encoded in H.264 at a gross bit rate of 2 Mbit/s including Ethernet overhead. It follows that this IPTV service requires 160 Mbit/s. The operator's IP core network will transmit the calculated 160 Mbit/s of IPTV traffic to the DSLAMs (Digital Subscriber Line Access Multiplexers) regardless of how many subscribers may be at the receiving end. This is a fair amount of traffic, but certainly not enough to have a dramatic effect on the IP core network's throughput.

Again, the core network typically distributes all IPTV channels en masse without factoring any individual channel's current usage into the equation. This means opportunities for saving bandwidth are limited. The access network is another matter: IGMP (Internet Group Multicast Protocol) snooping in the aggregation switches and edge routers can certainly help conserve bandwidth. A passive IGMP snooping mechanism can spy on IGMP Query, Report and Leave (IGMP version 2) packets transferred between IP multicast routers/ switches and IP multicast hosts to ascertain IP multicast group membership. It checks IGMP packets on their way through, pinpoints group registration info, and configures the multicasting routine accordingly. Without IGMP snooping, multicast traffic is treated just like broadcast traffic - that is, it is forwarded to all ports. With IGMP snooping, a group's multicast traffic is forwarded only to ports allocated to members of that group. And added proxy reporting can help reduce the number of IGMP membership report messages.

How video-on-demand services affect the network

VoD imposes severe bandwidth constraints on the network because video signals are streamed to the user in unicast frames. This consumes an inordinate amount of resources. The overall transmission rate equals the sum of all concurrent video streams.

Take, for example, a VoD service destined for 10,000 IPTV subscribers. This rule of thumb is that the network must be able to handle VoD requests for 10% of these customers. So, according to this baseline figure for budgetary calculations, the IP core network must be able to deliver 1,000 movies, all at the same time. Encoded in H.264 format at 2 Mbit/s gross bit rate, this constitutes a traffic load of 2 Gbit/s.

Why reliability matters so much

Networks tasked to deliver IPTV services must be, above all, reliable. Two goals figure prominently in all ResIP design efforts – to prevent any single point of failure and to achieve failover in less than one second in both the IP core network and the access/ aggregation network. The IP core network may be based on an Interior Gateway Protocol (IGP) such as IS-IS (Intermediate System to Intermediate System) or OSPF (Open Shortest Path First) or on an MPLS (Multiprotocol Label Switching) backbone. The access and aggregation network has been based on Ethernet, and in a recent implementation also on VPLS (Virtual Private LAN Service), with redundant interconnection of the video head-end equipment to the IP core network. Figure 1 shows an example of a resilient multicast network design.

This architecture calls for multicast replication in the DSLAM, and for multicast traffic to be distributed from the redundant IP edge to the DSLAMs via a multicast VLAN (Virtual LAN) or VPLS entity. Usually there is just one IPTV edge location in the network. Some situations require geographic redundancy, in which case there will be two IPTV edge locations. However, the customary configuration comprises multiple IP edges, one for each regional access and aggregation network. Several options come to mind for the access and aggregation network. If an engineer opts for VPLS in aggregation networks, he can stack a logical topology of paths and VPLS entities on top of any physical topology. What's more, services can benefit from the underlying MPLS network's features such as traffic engineering, fast recovery, and resilient multi-homing for the DSLAM.

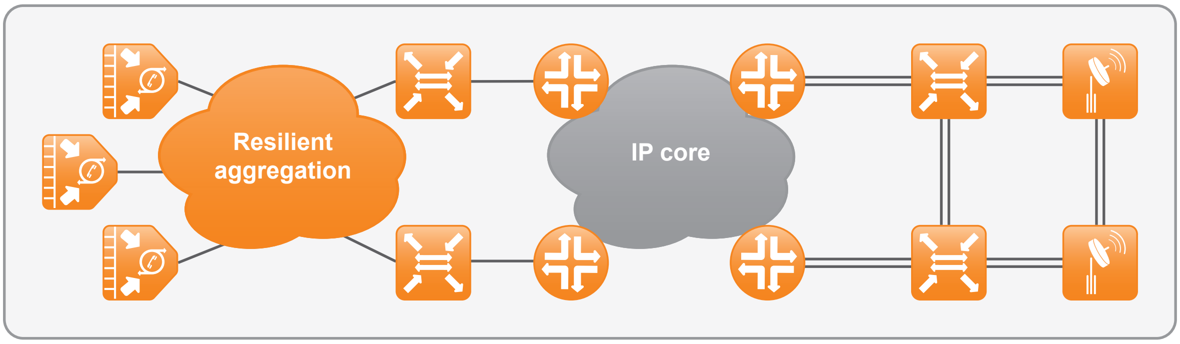

Figure 1: Redundant multicast network design

A high hurdle: quality of service
IPTV is a real-time service with exceedingly stringent QoS requirements, particularly when it comes to packet loss and delay. A little delay will not adversely affect the quality of experience. However, anything longer than one second may result in a less than a gratifying experience - especially if a neighbor is watching the same big soccer match on a conventional TV and starts celebrating a goal several seconds before the IPTV subscriber sees it scored. Packet loss is likely to cause visible artifacts due to the high compression rates of MPEG-encoded TV signals. The rate has to be lower than 10^{-6} to ensure less than one visible artifact per movie on the TV screen.

Quality of service is such a complex and important issue in network design that Nokia Siemens Networks has devoted a whitepaper to the subject. To learn more about the ResIP Center's QoS recommendations for all services including IPTV, please read the paper entitled **QoS-enabled IP networks**.

Zapping channels – frustration or elation?
Zapping speed – that is, how fast the viewer can change IPTV channels – is another measure of how rewarding the experience will be. Channel change delay is the time it takes for the first video frame to appear on the screen after the user punches a button on the remote control. The following factors affect this latency:

- The time it takes to process the remote command
- IGMP control processing for the set-top box (STB), residential gateway, and network
- Data plane protocol stack processing, including DSL interleaving delay, decryption, FEC, and the like
- STB jitter buffer delay
- Video encoding buffer
- MPEG decoder delay for re-synchronizing with the new program

Latency is also inherent in IGMP processing throughout the network, from the residential gateway to the DSLAM and BSR (Broadband Service Router), which lengthens delay time when zapping.

The jitter buffer must be sized according to the amount of jitter in the media stream received by the STB. A well-engineered, high-bandwidth IP network generates less than 50 ms of jitter. The video source should also produce streams with low jitter.

MPEG decoder latency generally accounts for the largest share of overall delay because the decoder usually has to wait for an I-frame to resynch. Beyond that, B-frames that are encoded using past and future I- and P-frames will add to the decoding delay. This delay depends on the compression rate, the encoding algorithm, and the number of B- and P-frames between two I-frames, which is called a group of pictures, or GOP for short. The compression rate certainly affects channel change delay: High compression rates entail a large GOP with many B-frames, and therefore cause far greater encoding/decoding delay.

When a user presses a button on the remote to change channels, the STB responds immediately by presenting a dialog box with the program name, time, channel, and so forth. This shows the viewer that the device is busy catering to his needs and spares him the sight of a dark screen while waiting.

The DSL Forum finalized the TR-101 specification for Ethernet-based DSL aggregation networks' fundamental architecture. TR-101 also encompasses definitions for the delivery of multicast services.

A look at access network architecture

TR-101 distinguishes between two modes of connecting broadband DSL users to the aggregation network. The first is called 1:1, also known as VLAN per subscriber. The second is N:1, sometimes called VLAN per service. More and more modern-day aggregation networks are MPLS-based, particularly when it comes to second-mile aggregation. Consequently, VPLS is often used to separate different traffic flows logically. VPLS-based aggregation also fits the architectural specifications of TR-101.

1:1 in action
1:1 mode focuses on the subscriber. It provides a simple connection to the IP network, and enables multiple service delivery. Introducing a new service does not require renewed VLAN configuration. Scalability concerns dictate the use of stacked VLANs.

The inner VLAN tag identifies the subscriber's port at the DSLAM. The outer VLAN tag identifies the DSLAM, and can be added by the DSLAM or by the next aggregation switch. Usually this mode is paired with a single edge router. In other words, a single BSR provides all services. In cases where a VLPS (Virtual Private LAN Service) serves aggregation purposes, the outer tag is mapped to a VPLS in this aggregation network. Accordingly, one VPLS entity is in place for each DSLAM. The inner tags are handed over transparently.

This model calls for a dedicated multicast VLAN alongside the subscriber VLANs. It is a service VLAN with a single tag. The multicast VLAN distributes multicast traffic from the IP edge to the DSLAMs, and transports IGMP messages between the multicast router and multicast hosts (i.e. the STBs). When VLPS serves aggregation purposes, this multicast VLAN is mapped to a dedicated VPLS service shared by all DSLAMs.

This multicast VLAN may be terminated at the BSR, or optionally at any other IP edge router. Both the DSLAM and aggregation switches support IGMP so that multicast distribution may be optimized on the fly.

1:N in action
Based on service-specific VLANs or VPLS services in the aggregation network, this alternative aims to improve service delivery through the access and aggregation network, and employ optimized edge routers to the IP core (multiple edge). 1:N entails a multicast VLAN, as well as specific VLANs for data or voice services. The multicast VLAN always comes in the guise of a service VLAN. In the case of VPLS-based aggregation, each of these VLANs is mapped to a VPLS service, so three VPLS entities are devoted to each DSLAM, one each for Internet, voice over IP, and video services. Beyond that, all DSLAMs share one multicast VPLS service.

1:N is usually tied to service-specific PVCs (Permanent Virtual Circuits) or VLANs on the DSL link. The residential gateway must forward the Ethernet frames via the appropriate PVC or VLAN service. These two options may be mixed and matched in the same network. For example, all service VLANs may be connected to a single edge router, or multiple C-VLANs (Customer VLANs) per subscriber may be connected to service-specific edge routers.

The benefits of Layer 2 control
A Layer 2 control mechanism can enhance both 1:1 and 1:N modes. The DSL Forum's TR-059 technical specifications defined queuing and scheduling mechanisms to avoid congestion in the access network while dealing with multiple flows with distinct QoS requirements. Commonly called hierarchical scheduling, this requires that the BSR has knowledge of the access network, the various links in use, and their respective rates. However, some of this information such as the DSL sync rate is dynamic. A provisioning or inventory management OSS (operation services support) system cannot provide such dynamic data on the fly.

However, a Layer 2 control mechanism can provide this dynamic information from the DSLAM to the BSR when a DSL line resynchronizes. Unlike with other proposed mechanisms using an extended DHCP (Dynamic Host Configuration Protocol) relay agent or PPPoE (Point-to-Point Protocol over Ethernet) intermediate agent, this information is available in the BSR irrespective of an active user session.

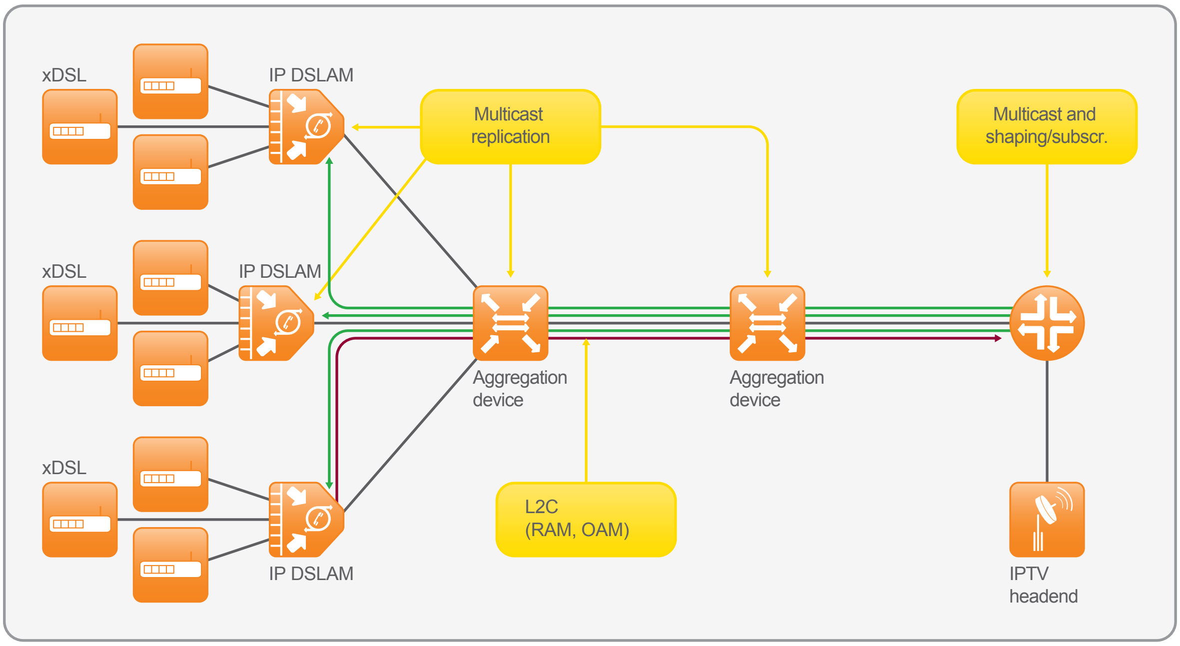

Figure 2: Layer 2 control and multicast replication

Beyond merely providing QoS support, Layer 2 control can also close the functional gap for end-to-end connectivity tests between the BSR and the CPE (Customer Premises Equipment). In the past, ATM's F4/F5 loopback tests served operating and troubleshooting purposes. Now Ethernet is figuring ever more prominently as a Layer 2 aggregation network, and as such must be able to test and troubleshoot connectivity in access networks with different technologies, to include the local loop.

But Ethernet OAM standards such as IEEE 802.1ag are still in the works, especially when it comes to mixed Ethernet and ATM setups. Layer 2 control, however, has bridged this divide and has proven its ability to test connectivity between the BSR and the CPE, including the local loop.

Replicate to elevate performance

The key to delivering IPTV broadcast services efficiently is multicast replication. As pictured in figure 2, all network elements - DSLAMs, aggregation switches, IP edge routers – can replicate multicasts, depending on network topology and service subscription rates. VDSL is a short-range link, so fewer subscribers are connected to each DSLAM in VDSL networks. In this case, the last multicast replication point can also be the first aggregation switch rather than the DSLAM.

Routing multicast traffic through the IP backbone

The following section examines three mechanisms for distributing multicast traffic throughout the IP backbone. Two, PIM SSM and PIM SM, are based exclusively on IGP. The third, P2MP LSP, is based on MPLS multicast distribution.

What PIM SSM is and what it does
Figure 3 outlines how Protocol Independent Multicast Source-Specific Mode works. The edge router must know the IP address S of the multicast source providing multicast group G. The router issues a PIM Join request, addressing it directly and sending it via the shortest path to the multicast source. The message goes to a router that is already aware of multicast group G, and the SPT (shortest path tree) is extended to the requesting edge router. It has two ways of pinpointing the multicast source's IP address S:

a) The receiver employs IGMPv3 to request a certain multicast group address G, and includes the multicast source address S in every IGMP Join message. IGMP v1/2 lacks this function.
b) The edge router is configured so that it is 'aware' of the G-to-S mapping scheme. When an IGMPv1/2 or IGMPv3 Join message goes out without specifying the multicast source's IP address, the edge router references this mapping scheme to identify the right source for multicast group address G.

The most convenient option is to bring PIM-SSM into play in the core and IGMPv3 in the access network because this entails the least configuration effort. The IPTV application server sends multicast groups and their sources' mapping information directly to the STBs with no need to maintain any particular status in the router network. The only prerequisite is that the STB supports IGMPv3, which must be confirmed first.

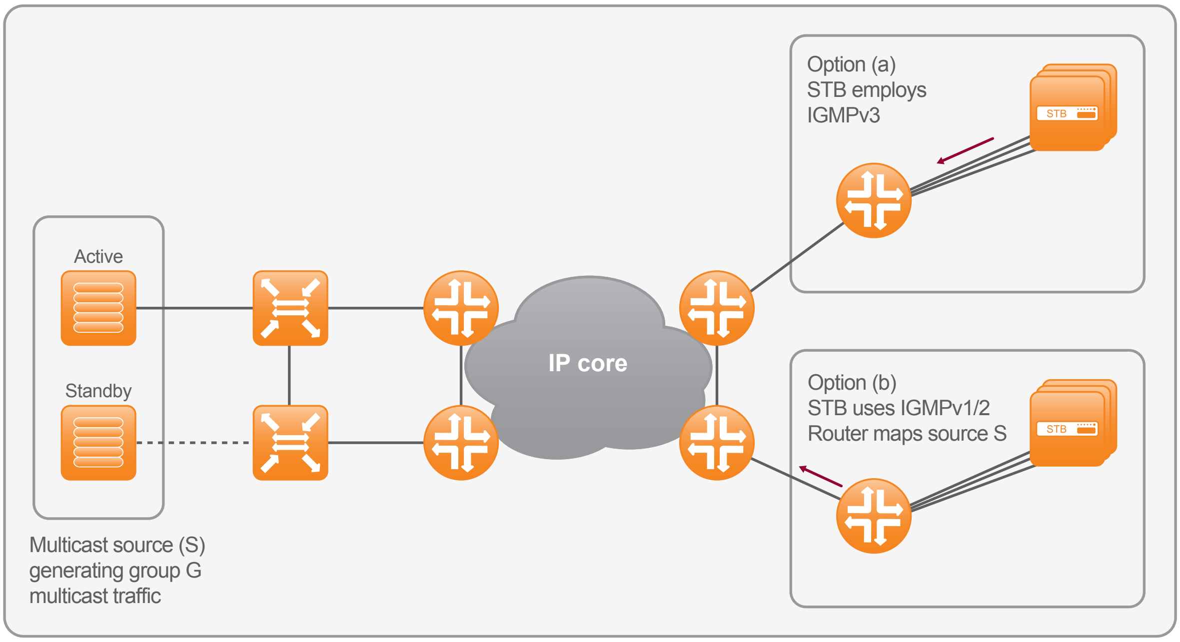

Figure 3: PIM SSM mapping scenarios

What PIM SM is and what it does

In contrast to PIM SSM described above, in Protocol Independent Multicast Sparse Mode, the edge router need not know which multicast source can provide multicast data for a specific group address G. The routers simply ask another router called a Rendezvous Point (RP) for this information.

This multicast distribution mode in the IP backbone is the way to go when there are many potential multicast group addresses, the multicast source IP addresses may not be known beforehand, or multicast source IP addresses change frequently.

There are several ways to choose an RP; the preferred is Anycast RP: Each potential RP router has a primary and a secondary loopback address. The secondary loopback address does not need to be unique throughout the network. The secondary IP address defines the RP, and each such secondary IP address must occur at least twice in the network. When an edge router is compelled to contact the RP, it is automatically patched to the nearest router, ascertained using IGP metrics, with a given secondary loopback address. If an RP fails, the IGP's convergence time determines how long a receiver is unable to join a new multicast group. Beyond that, the RP's load is balanced automatically.

In this context, PIM SM requires protocols and routing mechanisms that standard unicast traffic does not. Nokia Siemens Networks partners with both leading router vendors, Cisco Systems and Juniper Networks, and has developed a method of ensuring seamless interworking between these two vendors' routers, even in multicast architectures.

What P2MP LSPs are and what they do

The MPLS solution is less dynamic than the PIM solutions. The basic assumption here is that the edge routers always have active multicast receivers attached. All these edge routers are LSP egress routers of certain Point-to-Multipoint Label Switched Paths. Conversely, the ingress point of each P2MP LSP is a border router connected directly to a multicast source, in most cases via a redundant Layer 2 network. The configuration dictates that at this ingress point, incoming multicast data traffic is mapped to the corresponding P2MP LSPs.

When the source sends multicast data traffic, it is switched along the P2MP LSP and replicated wherever necessary. This multicast data traffic is then available at each P2MP LSP egress point, and will be forwarded to the subscriber in response to an incoming PIM or IGMP request.

The Resource Reservation Protocol serves to set up P2MP LSPs. Consequently, full traffic engineering capabilities are available for P2MP LSPs if the LSP is set up much like common point-to-point LSPs. Fast reroute may also be employed on the fly to protect against failures.

The bottom line

Again, PIM SM is the way to go in the IP backbone when there are many potential multicast group addresses, the multicast source IP addresses may not be known beforehand, or multicast source IP addresses change frequently. Accordingly, PIM SM is suitable for typical multicast applications such as voice, video conferencing (if delivered via multicast), and gaming. The ResIP Center recommends PIM SSM and P2MP LSPs for IPTV services.

Sub-second failover times can be achieved for physical and also Layer 3 failures by combining multicast features with intelligent implementations in routers and switches. Link failures in the core, router crashes, multicast source malfunctions - a concept encompassing both failure detection and recovery has been developed to ensure sub-second failover times for all these failure cases.

Credible concept, satisfying multicast service

Nokia Siemens Networks has a multicast solution that supports IPTV applications. Proofed and certified in the ResIP Center, its underlying concept is an integral part of access, aggregation, and core networks serving any fixed and mobile residential and business customers. Featuring well-designed subscriber management capabilities with zero-touch provisioning, this solution drives down operating costs.

Video services are far more unforgiving than voice when it comes to packet loss and delay, but users expect a great deal from services delivered over a broadband network. Nothing less than the quality provided by legacy video systems will do. Subscribers will not stand for interruptions and poor QoS brought on by a failure in the network.

Nokia Siemens Networks has developed a mechanism that maintains QoS by identifying various traffic types, and managing each according to its specific requirements across multiple links and network elements. This multicast solution also applies common security policies across the entire network to protect voice and video services against network-based and denial-of-service attacks.

The company's Consulting Service provides all the help customers need to integrate new products and applications in complex and heterogeneous multi-vendor networks, including customizing services. Network operators are welcome to participate in the ResIP Center's efforts and benefit from its outcomes.

Nokia Siemens Networks solutioneers have the insight necessary to gain an in-depth understanding of each customer's needs, priorities, and requirements, and the skill set to develop an optimized solution. Experts analyze the business requirements and tailor products and applications to the customer's unique demands. Before rolling out any solution, the ResIP Center and a team of globetrotting network engineers mitigate risk by subjecting each solution to a battery of performance and conformance tests, interoperability checks, and technical verifications to validate end-to-end functionality.

Living up to a persuasive performance promise, Nokia Siemens Networks delivers the right quality, at the right cost, and at the right time to accelerate integration and revenue generation while containing the customer's expenses. Any customer or prospect wishing to see the standard solution in action is welcome to schedule a trial on site or at one of the Nokia Siemens Networks Integration Laboratories.

Nokia Siemens Networks
P.O. Box 1
FI-02022 NOKIA SIEMENS NETWORKS
Finland
Visiting address:
Karaportti 3, ESPOO, Finland

Switchboard +358 71 400 4000 (Finland)
Switchboard +49 89 5159 01 (Germany)

Order-No. C401-00724-WP-201107-1-EN

www.nokiasiemensnetworks.com