

Achieving fast restoration times in IP networks for IPTV video transport – case study

Anna Wielosz, Kashif Islam, Cisco Systems, Inc.

Abstract: Recent deployments of IP based TV services have increased the importance of fast restoration times for service providers. For IP networks, fast service restoration requires fast convergence. The reliability of the IP network is compared to the TDM based networks used traditionally for video transport. This paper discusses network layers and protocols involved in the IP TV networks and how to achieve fast convergence for each of the contributing protocols. The proposed approach focuses on tuning the control plane parameters of the IP network with traditional routing protocols to provide the best possible restoration times. Finally, the convergence times' examples, with and without fast convergence improvements, for Broadcast video and VoD applications over the IP network are presented. The examples show that the convergence times can be reduced to sub-second periods compared to multiple seconds' outage in default configurations.

Introduction. Service Providers are increasingly using IP-based multi-service, converged backbones to support traditional and new, innovative services. As such, the resiliency requirements for those networks are becoming similar to traditional TDM networks. Traditional voice based SP's have long strived for and achieved network availability in the ranges of 99.999% for voice services. This is equivalent to 5 minutes of unplanned outage per year. This reliability expectation is now being applied broadly to converged IP networks.

Convergence times of 1-2 seconds were adequate for applications using TCP and designed to work in a bursty fashion with retransmission of missing packets as required. For those applications, impact of 1-2 second outage would have minimal impact to subscribers, often un-noticed.

The IP based video requirements are different for several reasons:

- High level compression means the loss of one packet can result in visible artifacts. Large MPEG GOP sizes means that under certain scenarios, packet loss can lead to a several second loss of video.

- The human eye has been trained over the years to be acute detector of poor video quality.
- Traditional video services are of very high quality and reliability, creating a high standard for IP based video SPs to meet.

As a result, achieving fast convergence times is becoming a top priority for SP's striving to offer or preparing to offer competitive video services.

	Bandwidth Up kbps	Bandwidth down kbps	Jitter	Latency	BER	Availability
Data	100	5000/300	High	High	10-4	Low
Voice	100	100	Med	Low	10-5	Very High
Broadcast (MPEG2)	5	4000 Std 15000 HD	Low	Med	10-10	High
VOD (MPEG2)	5	4000 Std 15000 HD	Low	Med	10-10	Med
Gaming	20	100	Low	Low	10-4	Low

For our purposes, we define the network outage event as an interruption in the traffic flow between the application source, e.g. video encoder, and the application receiver, e.g. set-top box. The length of the network outage is closely reflected (but not necessarily equal) by the service outage, which can be defined as a service interruption perceived by the service subscriber. It is important to note that there are other network impairments that can degrade video quality as perceived by the end user that are not related to a link or node outage event.

There are two components contributing to the overall restoration times: control plane and forwarding plane convergence.

Control Plane Convergence

Control plane convergence is the process by which routers use specific protocols to exchange link and device information which is needed update their routing table when a network topology change occurs. Convergence is complete when all updates are complete and routing tables stabilize and start reflecting the updated topology.

The network design challenge is to work with existing protocol standards and tune protocol attributes to allow very fast propagation of failure information. An example of such attribute may be the PIM Hello mechanism frequency allowing fast discovery of a loss of adjacency. In summary, all states leading to the failure recovery shall be understood and each state parameter properly tuned.

Forwarding Plane Convergence

The forwarding plane is the process by which routers forward the converged routing information to the ingress and egress interfaces. The forwarding plane convergence is completed when the affected traffic flows are restored. Forwarding plane convergence time and process is implementation dependent.

The total convergence time after an outage event occurs is made up of contributions from both planes. This paper focus is on achieving fast convergence for the control plane. Although forwarding plane contribution may be significant, it is platform implementation dependent and is outside of scope of this paper.

IP outage impacts on video applications. New technologies such as ADSL2+, VDSL2, and fiber access roll-outs are enabling SP to deliver speeds of more than 20 Mbps per household depending on loop length. These higher speeds enable traditional wireline based telephone companies and Internet Service Providers (ISP) to envisage new residential services. Of the services, variants of broadcast television and Video on Demand (VOD) applications’ are significant and prominent. Transport over IP is deployed or envisaged by many wireline service providers.

Video service requires high levels of availability. The existing video broadcasters have established a highly available service and consumers have the same expectations for new video providers. And since video is central to household entertainment, outages are sensitive.

IP network vendors and operators have focused on methods to improve IP network reliability and add more effective and fast redundancy mechanisms. Sub-second convergence times are achievable, with direct relevance to the emerging IPTV carriers.

The outage scenarios can include one of the following:

- Link failure: due to a fiber cut of an interface failure, the physical connectivity between two network elements is lost.
- Network element failure: due to a hardware failure or control plane failure, the network element becomes temporarily unavailable.

The video signals are encapsulated in MPEG2 or MPEG4 format by encoder residing in head end. The MPEG frames are furthermore encapsulated in IP.

The transport layer may be UDP or RTP. Historically, multicast IP traffic is sent over UDP. The RTP brings capabilities like packet sequencing, time stamps and as such better troubleshooting and monitoring capabilities of video streams. The use of RTP for IPTV streaming is seriously envisaged by service providers.

Depending on the set-top-box (STB) implementation capabilities and video encoding parameters, the visual impact of the outage may vary. A sub-second outage may produce a TV screen freeze, with or without additional visual artifacts. Longer channel unavailability, usually upwards of a few seconds, causes the “channel unavailable” message to be displayed.

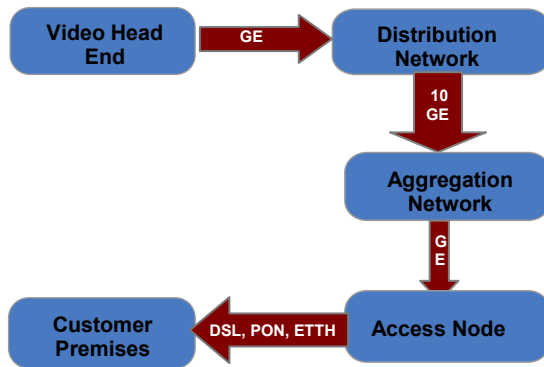
However, even very short outages, in the range of hundreds of milliseconds, are noticeable to the end user. Lab testing showed that link outage as short as 20-30 msec resulted in a noticeable artifact. Also, outages longer than 3 seconds resulted in channel unavailable. The table below summarizes a subjective classification of the visual outage impacts for the test bed. Please note that the “Channel unavailable” notification and time after which it is displayed will be vendor specific.

Light visual impact [s]	Noticeable visual impact [s]	“Channel unavailable” message [s]
0.03 to 1	1 to 3	3.5 and more

Pushing convergence times to even more stringent requirements does not eliminate a visible artifact, unless application layer error correction techniques are implemented.

The testing was performed to determine if the network topology based on IP routing protocols can meet targets for video restoration as discussed above.

IPTV network overview. The simplified IPTV network flowchart is depicted below.



The IP TV network will be composed of the following layers:

- Home Network (Customer Premises) – service provider (SP) and subscriber owned devices interconnected at the subscriber home. The central home device is designated as Residential Gateway (RG) and interfaces to the SP network. The RG interconnects home devices such as PCs, STBs and voice adaptors.
- Access – Subscribers’ local loops and the local loops’ aggregation devices. Today IP TV networks are mostly rolled over xDSL (ADSL2+, VDSL) and Ethernet technologies. The discussed deployment model assumes the access to be bridged layer 2 (L2) architecture for video services.
- Aggregation (Regional Network) – this network layer gathers traffic from many access devices and hands it off to the core or distribution network. The aggregation network is the layer 3 (L3) demarcation point for residential services in the model.
- Distribution (Service Provider) – this network layer provides high capacity and high availability transport. The distribution network can have multiple technologies enabled such as EoMPLS, VPLS, multicast. Usually it is shared between multiple services. This network layer is most often the injection point for video content. IP or IP/MPLS distribution is assumed for further discussion.
- Head ends – specific central offices (CO) where the video signal acquisition occurs and where most of the video equipment resides. The video equipment interfaces with the distribution routers where the video traffic is injected in either MPEG-2 or MPEG-4/H.264 over IP format.

Broadcast video. In our model, video is transported in native IPv4 format. The MPEG

encapsulation is not relevant to the convergence study and therefore unspecified. Although MPEG encoding parameters do influence how fast the stream can be displayed on the TV after connectivity restoration, it is not the subject of the study.

Broadcast video is transported using multicast IPv4. Currently most encoders are using UDP transport layer. RTP brings advantages such as capability of streams monitoring and is opening the door for error corrections techniques in the access. Next generation encoders most likely will be offering the RTP capability because of those advantages.

The video multicast sources are located at Video Head End. Each MPEG encoder encapsulates SPTS MPEG frames into IP with a unique multicast group address. The encoder is assigned with an IP address referenced further as the multicast source address.

The IGMP join issued from STB triggers a creation of the multicast tree.

If the STB issues the IGMPv2 join request, the request is converted to the IGMPv3 join which includes source address and multicast group (S,G). The aggregation router can perform this function, called SSM mapping.

When the IGMPv3 (S,G) join reaches the first router, the PIMv2 (S,G) join is created unless a particular group has already an active receiver on the aggregation router.

The PIMv2 (S,G) join is propagated upstream through the distribution routers towards the head end until the entire (S,G) multicast forwarding tree is built.

Video on Demand. VoD model uses unicast IPv4 transport. When a particular stream is requested, the STB initiates an RTSP session request to the VoD server. Once the session is successfully established, assuming the asset is present on the VOD, and the authorization has been granted, the VoD server starts forwarding unicast RTP or UDP stream to the STB. RTP in conjunction with RTSP is most widely used for VOD transport.

From this description, it is clear that the protocol stack involved in the convergence of VoD and broadcast video stream will be different. As well the route distribution mechanisms chosen for the two applications may (but not have to be) different. Therefore, the network outages resulting from the same failure may be different for both applications.

Network protocols for IPTV transport. With the network model as described in the previous section, the following protocols are likely to be used at each network layer:

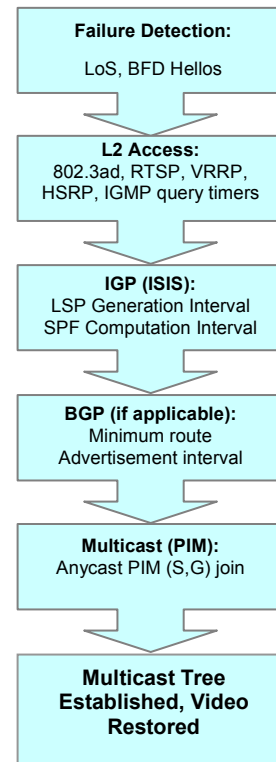
- Access: 802.3 or 802.1q Ethernet MAC based bridging. The DSLAM will perform bridging between DSL lines and the uplink. No IP routing is performed at the DSLAM. The DSLAM will broadcast multicast traffic over video VLAN unless the IGMP snooping mechanism is enabled.
- Aggregation and distribution: The Interior Gateway Protocol (IGP) is used to advertise network address reachability. ISIS or OSPF are most commonly used IGP protocols in distribution and metro networks. The PIM-SSM protocol is used for multicast routing.
- Video Head End: video head end addresses may be seen as internal or external addresses to the aggregation and distribution networks. In the case of internal addresses, the video head end subnets are advertised within the IGP. This can be achieved as the number of video prefixes is relatively low (today's implementations do not exceed 500 sources). Moreover, if the number of multicast source subnets is a cause of scalability concern, an external device can be used to aggregate multiple multicast streams to a single source.

If the video head end and distribution network are managed by separate operational entities or multicast content is provided by a different company, eBGP may be required. The multicast source routes are advertised in such a case via eBGP to the distribution routers. These routes can be further redistributed to the aggregation and core routers through iBGP or IGP. In such a case, the control plane of the network topology and services can be managed separately.

The PIM will build the (S,G) tree based on the unicast routing information from BGP or IGP, depending on how the multicast source reachability information is learned.

For the convergence discussion, the assumed model for the use case uses ISIS for internal routes' propagation and eBGP/iBGP for video prefixes. Such topology represents one of the worst cases when multiple protocol dependencies occur during the rerouting process triggered by a network outage.

Step by step convergence. When the network failure occurs, the following steps must be successfully completed for the full video service recovery.



Link failure detection. The efficient failure detection is the important step towards the fast recovery. Loss of Signal (LOS) or loss of adjacency, for which the discovery time can vary from microseconds to milliseconds range, triggers all related routing protocols to withdraw the unavailable prefixes, recompute new routes and advertise them.

Physical interfaces implement millisecond or even microsecond delay timers paired with signal thresholds in order to prevent declaring link down or up based on a transient condition. Secondly, the routers and switches may implement timers delaying reporting the failure to the higher layer. This type of timers is not standardized and it may not always be configurable (please refer to the conclusions for Cisco timers' settings examples).

The risk of setting the LOS timers too aggressively is causing excessive routing updates from a link. Link dampening feature should be used in conjunction with aggressive timers to prevent such a behavior.

Reducing LOS timers applies to direct links and physical failures. It does not help for other failures, such as router control plane failures or loss of router adjacency over non-direct links (DWDM with no LOS propagation, tunnels through MPLS, etc.). This is where the Bidirectional Forwarding Detection (reference [BFD]) may help in early detection of network failure..

As stated in [BFD]: *an increasingly important feature of networking equipment is the rapid detection of communication failures between adjacent systems, in order to more quickly establish alternative paths.* The BFD mechanism is a hello based protocol discovering loss of adjacency in the range of milliseconds. The negotiated *TxInterval* informs two adjacent nodes about the expected frequencies of hellos from each other. In asynchronous mode, when both nodes send control packet continuously, a failure is detected if no BFD packets have been received after *Detect Mult * RXinterval*. Assuming BFD *TXinterval* is equal to 50ms and the *DetectMult* (detect multiplier) to 3, a failure can be detected within 150ms.

This failure detection value is not as good as the physical LOS but much better than traditional hello based detection intervals (OSPF, ISIS, PIM).

Using BFD aware IGP and BFD aware PIM insures fast failure detection for those protocols and in turn enables fast convergence.

Fast hellos mechanisms can be alternatively used for the failures detection. In such a case, the IGP hello timers and PIM hello timers can be set to millisecond values. Default values for those protocols are summarized at the end of the paper.

Although reducing hello timers is possible, it is not as efficient as BFD. BFD will use one protocol instance per interface for all protocol adjacencies. In case of protocols' hellos, each protocol has its own hello per interface. The resulting overhead will be higher and the CPU utilization may be impacted.

Fast convergence on the access. In residential deployments, local loops are unprotected. The DSLAMs can aggregate hundreds of residential subscribers. A failure on a DSLAM uplink would cause impacts to all those customers.

The DSLAM uplink can implement the following protection scenarios:

- 802.3ad – link aggregation. Multiple GE links from the same DSLAM are terminated on the same aggregation router. The aggregation

router remains a single point of failure. Single link failure is protected with negligible traffic impact. There are no convergence delays, as the link scales down to the available bandwidth without any changes to control or forwarding plane. This approach requires bandwidth to be managed on the 802.3ad links in order to avoid congestion impacting video and voice under a failure scenario.

- Rapid Spanning Tree Protocol (RSTP) 802.1w or Multiple STP (MSTP) 802.3s – two uplinks are provided between a DSLAM and two separate aggregation routers. In the case of failure, the traffic converges towards the second aggregation router.

RSTP or MSPT provide fast convergence for bridged topologies. When the primary link fails, the bridge opens immediately on the standby link thereby achieving sub-second restoration.

Restoring L2 layer connectivity very quickly is not sufficient however to restore the video service.

To restore VoD unicast streams, the Virtual Router Redundancy Protocol [VRRP] (Cisco version is known as HSRP) can be used in conjunction with the RSTP. That allows the default gateway of the VoD application to move instantaneously from primary aggregation router to the standby aggregation router while keeping the same MAC address. In the VRRP protocol, the hellos are sent every 1 second between the routing instances. 3 lost hellos indicate to the standby instance that the master router is down. Therefore, the unicast convergence can be obtained in RSTP+VRRP time. Taking into consideration the VRRP timers, the unicast convergence in the range of 3 to 5 seconds can be obtained.

In the case of multicast, the situation is more complex. The standby link that has been brought up is not on the list of outgoing interfaces for any multicast group on the aggregation router. With the IGMP snooping on the DSLAM and on the aggregation router, the multicast traffic is forwarded only if the (S,G) receivers have been previously registered on a particular port. In this precise situation, the IGMP query has to be reissued on the switch where the topology change has been discovered [IGMP]. If the active receivers respond within the *MaxQueryTimeout* the IGMP join will be forwarded to the routed ports and the new link can be added to the outgoing interfaces of multicast groups that have receivers on the DSLAM. The DSLAM or

residential gateway are transparent and cannot reissue the IGMP join request.

As a result, the fast convergence in the access for the broadcast video may not be achievable in the L2 operation mode.

Fast convergence on the distribution/aggregation. The IGP convergence is the second step of the service restoration on the core, aggregation or distribution layers.

IGP. ISIS [ISIS] and OSPF are common IGP protocols used in medium and large backbones. This paper will use ISIS as an example but similar mechanisms to those discussed here are applicable to the OSPF.

ISIS routers exchange LSPs to build their routing information base. Each router has the view of the network routing topology. IGP is advertising the internal infrastructure addresses (including next hop BGP addresses).

Primarily, the key to achieve the fast convergence is the ability to propagate the failure information to all routers in the domain as quickly as possible. However, as ISIS uses LSP flood mechanisms to distribute information, under normal operation mode the timers prevent too frequent LSP generation. That protects from overwhelming routers with unnecessary routing updates that may cause high CPU usage and bandwidth use on the links.

In particular, the *minLSPGenerationInterval* and the *minLSPTransmissionInterval* are the important timers. The default LSP generation interval may be as high as 50ms which means this delay has to occur between the topology change discovery (e.g. LOS detected) and the LSP generation. If the timer is minimized to values as low as 1ms, it guarantees fast LSP propagation through the network. Possible drawback of lowering the timer for LSP flood is the excessive LSP generation in case of link flapping. The use of dampening, as explained in the previous sections, shall be used in conjunction with fast convergence timers.

The *minLSPTransmissionInterval* timer is preventing back-to-back LSP updates. Minimizing this interval under 500ms allows sending the update quickly if an update has been sent immediately before a failure.

Once the new LSPs are propagated, the Dijkstra algorithm is invoked on all routers to recalculate the SPF. The routers implement vendor specific timers that delay the SPF computations. This is done to avoid triggering the SPT computation

before the LSP propagation is completed which will in turn result in back-to-back SPT recalculation's process runs.

When setting this timer to low values for fast convergence, the time necessary to propagate LSP's from all routers has to be accounted for. Even a single failure may cause the LSP's generations from many routers therefore some buffer time shall be allowed. The timer shall be higher for networks with larger number of hops and with longer distances, where the propagation delays are considerable. Average number of hops and average propagation delays may be used to estimate the required timer.

In summary, providing short timers to allow fast LSP propagation in conjunction with fast trigger of SPT computation allow achieving sub-second convergence for ISIS.

BGP. BGP [BGPv4] is widely used in Internet infrastructures to advertise large amount of prefixes. BGP has not been designed for fast convergence but rather to exchange large amount of the reachability information and related metrics.

This section discusses the parameters required to converge quickly the routing information obtained via BGP.

Let's first consider the iBGP case. The regional network routers will obtain video source prefixes via iBGP sessions. The default behavior for iBGP is not to override the next hop information. Therefore, the border routers interfacing with video head ends will be advertised as next hop routers. The next hop reachability in turn is advertised through the ISIS.

Under a failure scenario not involving the border routers facing video head end, and assuming the reachability to those routers can be restored, the BGP information may remain unchanged. The only information that changes is how to reach next hop and this information has already been propagated through the ISIS. In this case, the BGP is not involved in the convergence mechanism.

If any of the iBGP routers loses the reachability to the next hop of advertised multicast sources, those routes must be immediately withdrawn.

By default, the iBGP routers trigger an immediate update with an explicit withdrawal of unfeasible routes. However, eBGP neighbors have to wait for the *MinRouteAdvertisementInterval* before they can issue a back-to-back advertisement. While this behavior is very important in Internet peering and transit networks, the IPTV eBGP sessions may

carry low amount of prefixes and both parts may agree to reduce this timer to 0.

Although the BGP standard does not specify such timers, additional holding timers to delay updates after an IGP topology change may be present. They would prevent excessive updates in case of flapping. Those timers can also be set to zero, assuming dampening is enabled on the links.

In summary, BGP impacts on convergence in IPTV networks would be most often negligible, assuming that the local RIB changes trigger BGP updates on the router immediately. For the eBGP neighbors, the *MinRouteAdvertisementInterval* should be changed to a minimal possible value.

The VoD convergence times presented in the test results section illustrate combined ISIS and BGP convergence times obtained on Cisco routers.

PIM-SSM. After the IGP and BGP fast convergence has been achieved, the multicast trees have to be rebuilt.

Before a failure, the PIM [PIM-SSM] adjacencies are built with all PIM neighbors. As soon as the IGP and/or BGP information has been updated, the new (S,G) tree creation is triggered.

All routers on which the reachability towards a source S has changed, must rebuild the (S,G) tree. The trigger will be either S RPF check failure or the PIM adjacency loss on the previously active link.

The router has to issue the (S,G) join towards the alternate neighbor if the PIM adjacency is already active.

PIM specification does not specify any timers for triggering the (S,G) joins upon topology change.

The broadcast TV test results in the *Test Results* section illustrate PIM-SSM convergence with ISIS and BGP underlying unicast protocols.

Fast convergence in head end. The head end multicast streams protection uses the mechanism called Anycast source. Anycast is a network addressing and routing scheme where traffic is routed to the destination with the best possible route as viewed by the routing topology. What it means in the IPTV environment, there are two or more encoders streaming the same content into the network with the same source IP address.

As PIM uses the underlying unicast routing table to build the (S,G) trees, multiple (S,G) trees will be built towards the shortest path sources.

In the case when one of the sources fails, the mechanisms described in previous sections are triggered. The head-end router withdraws the route to the failed source. All routers update the reachability information for the source IP address towards the alternative source which becomes the best path. PIM (S,G) tree can converge towards the alternate source. No additional timer or tuning is required.

Current challenge is how to inform the local head end routers about a source failure. In the case when a failure happens on a stream prior to the encoder, the IP interface of the encoder itself may remain active and the routers may still have the (S,G) attached to the wrong interface.

In other cases, an encoder or encryption engine, may handle multiple SPTS MPEG streams with the same source address. If one of the streams is experiencing problems, others are still active. Therefore, the IP source and all (S,G) trees remain active. Eventually, if there is no traffic coming for the particular (S,G) the tree will age out. The default setting is however 180s.

Currently, only vendor proprietary workarounds exist to resolve those issues.

Failure recovery considerations. The failure recovery will benefit from the previously described mechanisms. The unicast traffic will not be forwarded towards the new path before the control plane is completely restored. Therefore, no or very low packet loss shall be expected in such scenario.

The multicast traffic restoration has more components and the order of the restoration is crucial:

1. Restoring PIM neighbor adjacency.
2. Restoring best route to the source S.
3. Prune (S,G) from alternate route.
4. Join (S,G) to the restored best route.

The steps 3 and 4 cannot happen before the step 1 and 2 is completed. However, if the step 1 has not been completed before the step 2 is executed, it leads to the outage that may take up to multiple seconds.

As the unicast route has been restored, it results in the update of the next hop information for the source S. As a result, the RPF check will fail and the Prune (S,G) message will be sent to the current PIM neighbor. However, if the PIM adjacency on the restored link has not been built yet, the Join

(S,G) message will not be issued. That results in the loss of the multicast traffic for the group.

A delayed join will be issued once the PIM adjacency is established and after the *t_override_default* is expired (randomized delay between 0 and 2.5 seconds). If the first join was unsuccessful, the *t_periodic* timer has to expire, which is 60 seconds by default.

Instead of modifying join timers, it is more suitable to insure that the PIM adjacencies are established as fast as possible. If the problem is observed with unicast converging faster than PIM, reducing hello timer or adjusting *Triggered_Hello_Delay* on PIM neighbors may insure the proper PIM adjacency establishment.

Forwarding plane considerations. The forwarding plane comes into play after the control plane convergence is completed. Once the router has received all the information from the control plane, the information has to be propagated to the line cards. Only at that point the actual traffic flows can be restored.

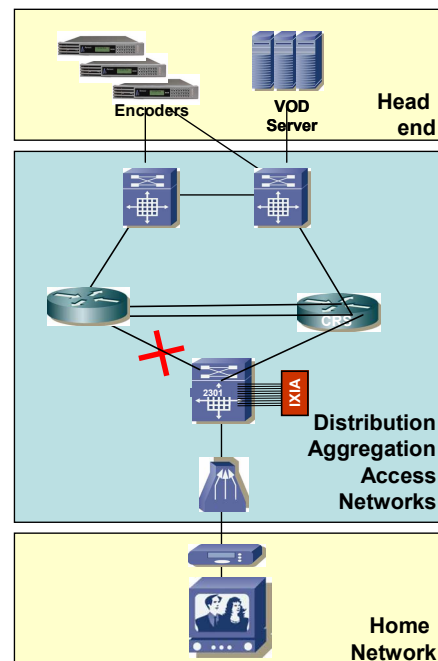
In heavy loaded environments, multiple factors can affect the forwarding plane convergence: the number of prefixes in the network, number of multicast groups that must be updated, number of MPLS labels in use, etc. These impacts are platform dependent and the best way of characterizing them is via conducting the performance testing with the real environment simulations.

Test results. – A typical network failure is a two stage process – failure and recovery. Convergence times have to be measured for both failure and recovery to get the complete picture of the network convergence performance. It is important for the network to return to pre-failure state once the failure has been corrected.

As stated earlier in this document the natural order of things is for the unicast traffic to converge before any multicast traffic. Because of this, VoD convergence is expected to be shorter than broadcast TV.

The test results included in this section are an example of a network carrying multiple residential services. The results are not intended to study the impact of each individual timers or mechanisms discussed in the paper. The tests are intended to illustrate the convergence improvement that may be achieved if some of those mechanisms are applied.

Test topology - The following diagram illustrates test network topology and with the indication of the links on which failures were simulated.



The network scenario was simulating service provider network with 400 multicast groups. In the background, residential and business services were also provisioned on the network. The traffic was present on all multicast groups and background services.

The failure scenario was simulating a link failure between aggregation and core routers. The control plane is assumed to be converged when the traffic starts being forwarded and is received by the end user through an alternate route.

The restoration scenario was measuring the time between the link reconnection and the traffic restoration. The restoration is assumed to be completed when the traffic starts being forwarded and is received by the end-user through the pre-failure route.

At first, link failures are created and measurements are taken with the default protocol timers. Then, the same failure scenarios are repeated with the fast convergence tweaks.

Multiple measurements were taken in each of the scenarios. The tables summarize observed outage times' ranges.

BTV convergence with 4K VCs and 400 channels		
	Default Timers	Fast Convergence
Failure Time [s]	2.5 – 4	0.3 – 0.7
Recovery Time [s]	0.1 – 0.2	0.005 – 0.1
VoD Convergence		
	Default Timers	Fast Convergence
Failure Time [s]	2.5 – 3.5	0.02 – 0.04
Recovery Time [s]	0	0

The results indicate that the fast convergence tweaks suggested in this paper make is possible to achieve sub-second convergence for Broadcast TV streams even in a scaled configuration.

Conclusions. The step by step convergence analyses show that the sub-second convergence can be successfully achieved in IP networks for unicast and multicast traffics. For service providers with IP networks deployed today, there is a possibility to engineer existing networks and achieve the convergence levels acceptable for residential services such as video.

When designing the networks with fast convergence requirements, each protocol should be analyzed and if required, the timers should be adjusted. The following table summarizes the timers per protocol with their suggested default values and proposed modifications.

Timer values	Default	Fast convergence
Physical		
LOS carrier delay	10ms (GE)* 2s (POS)	0
BFD	Not specified	50 ms **
Hello Intervals	30 s (PIM) 3s (ISIS)* 10s (OSPF)*	1 s **
IGP (ISIS)		
Min LSP Transmission Interval	50 ms * 5 s	1 ms
SPF computation delay	5.5 s *	100 ms
BGP		
Min Advertisement Delay	30 s	0

* Cisco default timers (Cisco 7600 router)

** Not applicable to the test bed. Minimum values are dependent on platform capabilities.

Another important fact to notice is, what observing video restoration brings to evidence, that even short impacts may result in visible artifacts. Pushing convergence time requirements to very short times does not eliminate a visible artifact, unless application layer error correction techniques are implemented.

References

[BGPv4] “A Border Gateway Protocol 4 (BGP-4)”, RFC 1771, March 1995

[BFD] “Bidirectional Forwarding Detection”, draft-katz-ward-bfd-01.txt, IETF, August 2003

[ISIS] “Information technology — Telecommunications and information exchange between systems — Intermediate System to Intermediate System intra-domain routing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service”, ISO 8473 November 2002

[PIM-SM] “Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)”, RFC 4601, August 2006

[IGMPv3] “Internet Group Management Protocol, Version 3”, RFC3376, October 2002

[VRRP] “Virtual Router Redundancy Protocol”, S. Knight, et al., RFC 2338, April 1998

Anna Wielosz – Network Consulting Engineer in Advanced Services IPTV practice team. Anna Wielosz joined Cisco Systems last year bringing 8 years of technical experience in the service provider area. As Associate Director - IP Access at Bell Canada and as independent network consultant, Anna has been involved in various projects implementing leading edge technologies in the telecommunications service provider environment, including IPTV developments. Anna has also been representing Bell Canada at MPLS and DSL Forums. Masters Degree in Telecommunications Network Management, Ecole Franco-Polonaise, Poznan, Poland.

Kashif Islam - Software/QA Engineer in Network System Integration Test Engineering team. Kashif Islam has been working with Cisco Systems for more than 6 years now. Kashif Islam has been a part of Carrier Ethernet and Broadband Triple Play solution verification team for more than 4 years. In his role as a Software/QA Engineer, Kashif has worked on design and verification of various Carrier Ethernet and Broadband Triple Play

solutions for multiple service providers. Masters of Engineering (Internetworking) degree and certifications CCIE (RS), CCNP, CCDP, CCNA, CCDA