High Capacity Carrier Ethernet Transport Networks

A. RASMUSSEN, J. ZHANG, H. YU, R. FU, S. RUEPP, H. WESSING, M. BERGER
DTU Fotonik
Technical University of Denmark
DTU Bldg. 343, DK-2800 Kgs. Lyngby
DENMARK
msbe@fotonik.dtu.dk   http://www.fotonik.dtu.dk

Abstract: - Ethernet as a transport technology has, up to now, lacked the features such as network layer architecture, customer separation and manageability that carriers require for wide-scale deployment. However, with the advent of PBB-TE and T-MPLS, it is now possible to use Ethernet as a transport technology, making the use of Ethernet as a convergence layer for Next Generation Networks a distinct possibility. Triple Play services, in particular IPTV, are expected to be a main drivers for carrier Ethernet, however, a number of challenges must be addressed including QoS enabled control plane, enhanced OAM functions and survivability. This article provides an overview of PBB-TE and T-MPLS and demonstrates how IPTV services can be realized in the framework of Carrier Ethernet.

Key-Words: - Carrier Ethernet, IPTV, T-MPLS, PBB-TE, NGN

1 Introduction

For many years, the telecommunication service providers have been looking for new ways to compensate the declining revenue on the plain old telephone service and broadband access. The next big step that can deliver a substantial increase in revenue is believed to be the introduction of IPTV. As a response to the increased interest in television via the Internet (IPTV) the Danish Advanced Technology Foundation decided to finance a new research project entitled High quality IP network for IPTV and VOIP (HIPT). The objective of the HIPT project is to enhance carrier Ethernet transport for IPTV applications by developing technology that can fulfill the increasing requirements in terms of bandwidth and quality and at the same time reduce cost of network operation.

Metro Ethernet forum (MEF) has provided a clear definition of carrier Ethernet in relation to Ethernet services. According to MEF, a Carrier Ethernet service and the underlying equipment supporting the service must possess the following attributes: Standardized services, Scalability, reliability, Quality of Service and Service management. This article provides in section 2 an overview of Carrier Ethernet transport technology with focus on Traffic Engineered Provider Backbone Transport (PBB-TE) and Transport-MPLS (T-MPLS). Following that, in section 3, the high level network architecture for IPTV transport over Carrier Ethernet adopted by the HIPT project is introduced. Finally section 4 lists the main research challenges envisaged by the HIPT project to make Carrier Ethernet comply with IPTV transport. The main areas for further research are identified and solutions as well as future directions are given.

2 Carrier Ethernet Technology

Telecom carriers have spent more than 10 years developing a Next Generation Network concept that will allow them to simultaneously deliver packet-based and circuit-based services. It is now widely accepted that the Internet Protocol (IP) will form the basis for new services, as well as assist in the transition of circuit-based services to packet-based services (e.g. Voice and Video over IP).

However, it is far from certain that IP routing technology will be adopted as the transport convergence layer. IP/MPLS has been widely deployed, especially in carrier backbone/core networks as a service layer and as a convergence layer, but Ethernet is fast becoming a credible alternative candidate.

95% of all data traffic either originates or terminates as Ethernet, and data volume is forecast to grow tremendously given the impact of new video services and IPTV, for example. This situation has prompted many telecom carriers to consider Ethernet as a potential convergence solution for Next Generation Networks. With its scalability, ubiquity and natural support for IP services, Ethernet provides a compelling case. But before Ethernet can be adopted, it must be capable of supporting multiple services with at least the same level of quality as existing carrier services. In other words Ethernet must achieve a carrier grade of quality.

2.1 Provider Backbone Transport

IEEE has developed a number of standards providing enhancements to the original Ethernet standards (dating from over 30 years ago). These include [1-5]:

  802.1Q: Virtual LAN

802.1ad: Provider Bridging
802.1ah: Provider Backbone Bridging (draft)
802.3ah: Ethernet in the First Mile (with OAM)
802.1ag: Connectivity Fault Management (OAM)

PBB-TE builds upon these standards to provide a network solution designed specifically for transport applications. PBB-TE creates an independent connection-oriented packet-switched transport layer. This allows various services, not necessarily limited to Ethernet services, to be transported transparently through the network.

The following sections will concentrate on key features, which form the foundation of PBB-TE operation.

2.1.1 Provider Bridging (PB)

802.1Q VLAN capabilities were originally developed to provide a logical division of the same physical Ethernet network. However, only a limited range of possible VLAN instances was defined (4094 to be precise). This is the so-called "Single Tag".

In order to improve scalability, equipment vendors added support for a second VLAN tag. The resulting "Q-in-Q" or "Double Tagging" mechanism has been formalized in the IEEE 802.1ad Provider Bridging revision to 802.1Q.

The inner tag field or C-Tag, see Figure 1, carries the customer VLAN Identifier (C-VID), which identifies a customer VLAN (C-VLAN). The outer tag field, or S-Tag, carries the S-VID, which identifies a service VLAN (S-VLAN). This tag is used to identify a service instance and defines a topological partition of the network based on the topology of this service instance. Spanning tree protocol is used to prevent loops in each S-VLAN (and independently, to prevent loops in each C-VLAN). S-VLAN provides customer separation and also isolation of customers from a carrier's network. However, the S-VLAN tag is itself too limited for large-scale carrier networks (ie. more than 4000 customers).

2.1.2 Independent VLAN Learning

As the name suggests, with IVL there is an independent MAC address table for each VLAN. Forwarding is based both on the VLAN and destination MAC address. With IVL the MAC address table is local to a specific VLAN, which allows isolation of addressing within a customer or service VLAN instance.

Double tagging can be used with Independent-VLAN-Learning (IVL), defined in 802.1Q. However, the tag space limits the number of S-VLAN to 4094, which limits the number of service instances that can be deployed in a single network. Thus, Ethernet's scalability problem remains.

A consequence of partitioning the single Ethernet layer network is that S-VLAN specific state must be maintained by all switches in the network, which further limits scalability. In addition, VLANs are assigned across all ports in a switch, so reuse of VLAN IDs on different links is not possible.

2.1.3 Provider Backbone Bridging (PBB)

PBB (also known as MAC-in-MAC) encapsulation adds layer-networking support to Ethernet. MAC-in-MAC encapsulation is now being formalized in the 802.1ah 'Provider Backbone Bridges' draft standard.

Client PB Ethernet frames are encapsulated and forwarded in the backbone network based on new B-DA, B-SA and B-VID backbone-destination-address, backbone-source-address, and backbone-VLAN-ID fields.

MAC-in-MAC encapsulation support improves upon the separation and isolation features introduced in 802.1ad: it supports complete isolation of individual client-addressing fields as well as isolation from address fields used in the operator's backbone. 802.1ah also introduces a new 24 bit tag field; the I-SID service instance identifier. This 24-bit tag field is proposed as a solution to the scalability limitations encountered with the 12 bit S-VID defined in Provider Bridges.

802.1ah Provider Backbone Bridges operate the same way as traditional Ethernet bridges. Service is still connectionless, flooding is used when destination MAC addresses are not recognized, and spanning tree is used to prevent loops. VLAN tags are reserved on a network, rather than a per-port basis.
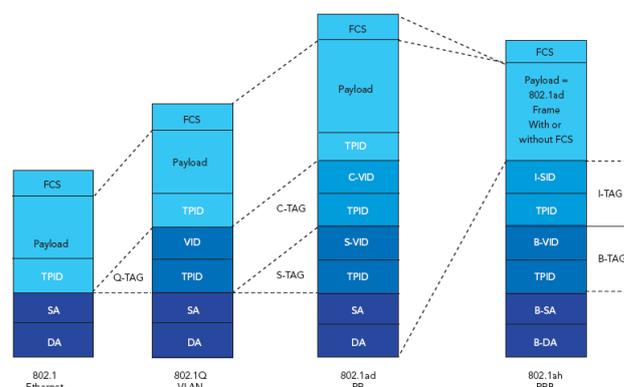

Figure 1   Ethernet frame format and evolution to Carrier Ethernet

802.1aq offers shortest path bridging, which uses a different set of shortest path trees (one for each VLAN) on each node, rooted at that node, as an improvement to Spanning Tree Protocol (STP). Rooting the tree on a per-node basis is required to ensure a shortest path from one node to any other node. The objective of shortest path bridging is to guarantee loop free forwarding, while using shortest path routes. 802.1aq proposes to use IS-IS protocol to update link state information. While the initiatives above have proven to be extremely useful in solving issues with

Ethernet as a carrier technology, there is growing concern that Ethernet itself is now becoming too complex, which compromises some of the key advantages of Ethernet, namely familiarity, simplicity and ease of use.

The proponents of PBB-TE have provided an alternative operational model based on simplification and focus. Existing technology is re-used to as great an extent as possible, while enhancements are focused on the specific needs of carrier transport networks. In other words, PBB-TE is the application of Ethernet to transport networks.

2.1.4 Key Features of PBB-TE

Proposals for PBB-TE are based on existing Ethernet technologies, including:
802.1Q VLAN
802.1ad PB
Subset of 802.1ah PBB
Subset of 802.1ag CFM

PBB-TE reuses existing implementations of VLANs and double tagging and combines them with the network separation and layering principles of PBB.

However, it does not use flooding/broadcasting mechanisms or Spanning Tree protocols - these are switched off by definition - and is thus intended to be used in connection-oriented network applications.

In this way, PBB-TE can be operated in a managed, deterministic connection-oriented fashion, similar to other existing transport technologies such as SDH/SONET.

PBB-TE operation can be clearly differentiated from PBB. In the PBB model, the B-VID identifies a packet flooding domain which interconnects different PB networks. In the PBB-TE model, the B-VID in combination with the B-DA address identifies a specific path through the network,.

The following sections outline these key characteristics in more detail.

2.1.5 Forwarding Model

PBB-TE is intended to be deployed as a connection-oriented packet-switched network layer. PBB-TE exploits the forwarding-engine of Independent VLAN Learning (IVL) capable switches, which allows packets to be forwarded based on the 60-bit concatenation of the B-VID and Destination MAC address. For PBB-TE, MAC learning functions are disabled, which means that the forwarding table's B-VID+MAC entries must be set by cross connect software on instruction from a management system or a suitable control plane. In PBB-TE, broadcast frames are discarded.

PBB-TE is intended to run on any IVL capable switch. An additional requirement that these switches need to meet is that packets with unknown VID+ MAC entries are not flooded but discarded. The switch must also support software configuration of VID and Destination MAC

forwarding table entries. Both 802.1Q virtual bridged and the relevant subset of 802.1ah provider backbone bridged operation should be supported.

Globally unique addresses simplify end-to-end network operation. Unique addresses reduce the potential for misconnection that can occur with technologies that use addresses with link or sub-network scope. The 12-bit VID field is used to identify alternate paths to the associated destination MAC address, e.g. to support protection switching.

2.2 Transport-MPLS

MPLS was originally developed by IETF in order to address core IP router performance issues, but has since found strong application in carriers' converged IP/MPLS core networks, and as a platform for data services such as IP-VPN.

With increasing packet networking, the ITU-T became interested in adapting MPLS to make it a "carrier class" network, functioning according to recognized ITU-T architectural principles. The result is Transport MPLS (T-MPLS), a connection-oriented packet transport network based on MPLS that provides managed point-to-point connections to different client layer networks (such as Ethernet).

However, unlike MPLS, it does not support a connectionless mode and is intended to be simpler in scope, less complex in operation and more easily managed. Layer 3 features have been eliminated and the control plane uses a minimum of IP - this should lead to equipment implementations that support carriers' needs for lower-cost, high-volume packet networking in their next-generation architectures.
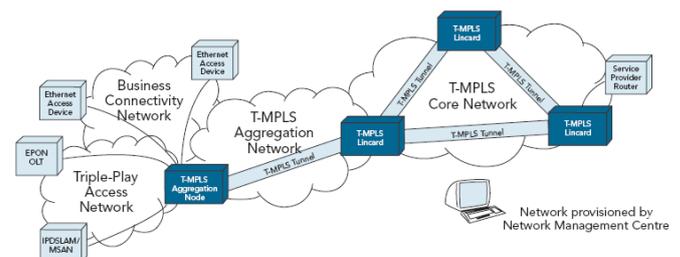

Figure 2 Example T-MPLS network architecture

T-MPLS is formulated in conjunction with today's circuit-based transport networks, following the same architectural, management and operational models. It is thus intended to provide an optimum evolution path for many carriers in their metro and access networks, as they transition to a packet-based future.

2.2.1 T-MPLS Standardization

ITU-T Study Group 15 has been standardizing the definition of T-MPLS since early 2005, with the stated objective of

identifying a subset of existing MPLS necessary and sufficient to provide connection oriented packet transport.

T-MPLS is initially focused on supporting Ethernet services but the client/server architecture can handle all packet services, including IP/MPLS, in concert with SDH/OTH circuit services.

One of the main definition goals of T-MPLS was to be aligned with ITU-T Recommendations G.805 and G.809 **Error! Reference source not found.Error! Reference source not found.**which themselves define a layered network architecture model for today's transport networks.

This standardization approach has been supported by a number of leading systems houses, with a combined market share of well over 70% of the WDM/SDH transport equipment market and indeed some early and target product platforms are already on the commercial market.

In 2006, ITU-T approved the first three Recommendations on Transport MPLS after various developments during the year in co-operation with IETF and MFA Forum. These three documents are:

• G.8110.1: Architecture of Transport MPLS (T-MPLS) Layer Network;

• G.8112: Interfaces for the Transport MPLS (T-MPLS) Hierarchy (TMH);

• G.8121: Characteristics of multi-protocol label switched (MPLS) equipment functional blocks.

T-MPLS scope is large and further standardization work is ongoing, including:

• New amendment to G.8110.1, due for consent in 2007;

• OAM requirements and mechanisms (G.8113, G.8114), also scheduled for consent in 2007;

• Linear and ring protection methodologies (G.8131,G.8132);

• IETF T-MPLS work under the moniker "draft-bryantpwe3-mpls-transport-00.txt" in the PWE3 WG

2.2.2 Definition of T-MPLS

T-MPLS might best be described as a transport network profile of IETF RFCs and G.8110 (the ITU-T's definition of MPLS), and is defined as a strictly connection-oriented subset of MPLS, meaning that:

• Forwarding behavior of T-MPLS is a subset of IETF defined MPLS. This common data/forwarding plane retains the essential nature of MPLS and ensures that interoperability and interworking will be readily achievable.

• Survivability is specific to the transport network. See section on Resiliency and survilability

• OAM is specific to the transport network and functionality is referenced from ITU-T's Y.1711 (OAM mechanism for MPLS networks).

• T-MPLS control plane (specific for the transport network) is currently null. In other words, the management plane will be used for manual/automated provisioning, in the same

way as SDH and OTN/WDM networks are provisioned today.

• No label reservation. That is, T-MPLS will not reserve labels for its own use independently of MPLS. Any requirements for special label assignment will be handled by IETF and co-ordinated with the MPLS standards. Again, this helps to ensure that interoperability and interworking will be readily achievable.

2.2.3 Differences from MPLS

In order to define a subset of MPLS that is connection-oriented and that can be readily dealt with using the established transport OAM model, several MPLS protocol features have been excluded from T-MPLS. Key differences of T-MPLS compared with MPLS include:

• Use of bi-directional LSPs (Label Switched Paths). Whilst MPLS LSPs are uni-directional, transport networks conventionally provision bi-directional connections. T-MPLS therefore pairs the forward

and backward LSPs to follow the same nodes and links.

• No PHP (Penultimate Hop Popping) option. PHP, by removing the MPLS label one node before the egress node, simplifies the egress processing required. Indeed, it comes from a historical legacy of wanting to minimize router processing requirements. However, the interface now has a mix of IP and MPLS packets and the final node must perform an IP (or other payload) look-up instead. More importantly, OAM is more complex or even impossible since the MPLS label context is lost.

• No LSP Merging option. LSP Merge means that all traffic forwarded along the same path to the same destination may use the same MPLS label. Whilst this may promote scalability, in fact it makes effective OAM and Performance Monitoring (PM) difficult or even impossible, since the traffic source becomes ambiguous and unknown. It is thus not a connection-oriented concept.

• No ECMP (Equal Cost Multiple Path) option. ECMP allows traffic within one LSP to be routed along multiple network paths. Not only does this require additional IP header processing, as well as MPLS label processing, but it makes OAM more complex since Continuity Check (CC) and PM flows may follow different paths. This concept is not needed in a connection-oriented network.

3 IPTV over Carrier Ethernet

Layer 3 and IP have proven useful in addressing Internet and other best effort data applications, however, this approach is not well suited to high-bandwidth, critical services, such as IPTV, which, in general, cannot tolerate delays in the network. In the HIPT project, we intend to investigate whether more intelligent layer 2 and layer 1 networks can be used to alleviate problems anticipated in

current IPTV networks. Using PBB-TE/PBT and T-MPLS, autonomous network decision making is removed and more traffic engineering is performed. This ensures control over exactly where traffic is being transported in the network with the further ability to monitor individual traffic flows (which is not easy to accomplish in layer 3 networks).

This approach has the further advantage of being able to support layer 3. Rather than replace layer 3, the solution is intended to supplement it reducing the need for costly nodes supporting services. Thus, instead of deploying a large number of nodes with large complexity, a simpler, yet intelligent, layer 1/layer 2 network based on Carrier Ethernet Transport can reduce cost and complexity, while enabling independent scaling of IPTV services at layer 3 with fewer nodes.

The carrier Ethernet based network architecture for IPTV transport is based on a Layer 2 approach with Layer 3 support in the edge routers and L3 awareness in the DSL Access Multiplexer (DSLAM) as shown in Figure 3.
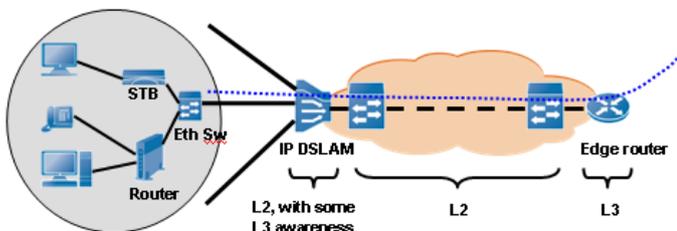

Figure 3: HIPT Network architecture

For simplicity, only DSLAM access is shown, but other types of last mile technology could be applied as well. IPTV flows are terminated in the STB (Set op Box) in the home network. The L2 network part between the IP DSLAM and the Edge router is assumed to be based on either T-MPLS or PBT-TE. In both cases, the objective is to transport IPTV signals with carrier class quality but at the same time reduce cost by utilising carrier Ethernet technology, which is assumed to be simpler than alternative layer 3 approaches, e.g. IP/MPLS.

Efficient transmission of IPTV traffic requires that the L2 Carrier Ethernet network is able to deliver sufficient capacity and Traffic management cababilities. The goal of the HIPT project is to develop high capacity PBB-TE and T-MPLS network nodes with Traffic management, OAM and QoS support to enable transport of demanding real-time applications, including IPTV. Furthermore, the objective of HIPT is to enhance the L2 network with a scalable control plane for traffic engeneering and resource reservation purposes. The next section will elaborate more on the research challenges involved in reaching these goals.

4 Research challenges

In the HIPT project, three main areas of research are described in the following sections.

4.1 QoS enabled control plane

For PBT and T-MPLS, a dynamic control plane is a key area for future development. Today, the technologies deliver managed end-to-end connections similar to SHD/SONET but new demanding applications in terms of quality and bandwidth would require control plane actions to guarantee proper behaviour of the application.

To develop dynamic a carrier Ethernet control plane, compatibility with The ETSI TISPAN RACS and ITU RACF specifications will provide a roadmap for Next generation Networks (NGN) compliance. The Resource and Admission Control Subsystem (RACS) layer is the TISPAN next-generation network subsystem that is responsible for elements of policing control, including resource reservation and admission control in the access and aggregation networks.

Similarly, the ITU Resource and Admission Control Functions (RACF) is an NGN control layer that supports dynamic reservation of resources and configuration of policy enforcement functions. In particular, the Access Transport Control Function (A-TRCF) could be mapped to an access bandwidth manager to enable resource and admission control in the Carrier Ethernet domain.

Figure 4 shows RACF in the NGN reference architecture based on Y.2111 that HIPT will adopt to provide a QoS enhanced control plane.
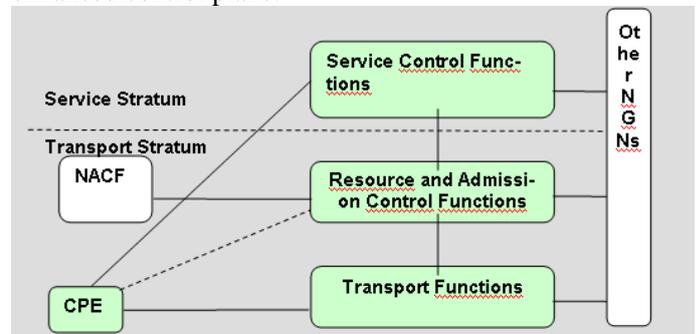

Figure 4 RACF in NGN

Carrier Ethernet resides in the Transport Functions block in Figure 4. Dynamic routing in the Metro/access domain will impose scalability limitations, so HIPT proposes a scalable architecture where static tunnels in PBT or T-MPLS are configured. The RACF will then be able to reserve bandwidth within the tunnels for a given application. This approach will simplify resource control and admission control and make it feasible to implement in the carrier Ethernet domain.

4.2 OAM for IPTV flow monitoring

With the IPTV service, the viewers will directly watch the quality of the network, thus it is very important that the network delivers carrier grade quality. Today, the networks do not deliver the required OAM functionalities and IPTV operators are forced to deploy very expensive solutions to monitor the TV signals. Typically, monitoring the quality is done by decoding all the TV signals and showing them on screens on various locations in the network.

Draft standard IEEE 802.1ag Connectivity Fault Management (CFM) has been recently developed to address the lack of end-to-end OAM in traditional Ethernet networks. It is closely aligned with the ITU's Y.1731 Recommendation (which also defines Performance Monitoring functions). Key functions such as loopback at specific MACs, linktrace to identify network paths and continuity check are defined.

OAM is specific to the transport network and functionality is referenced from ITU-T's Y.1711 (OAM mechanism for MPLS networks). This provides the same OAM concepts and methods (e.g. connectivity verification, alarm suppression, remote defect indication) already available in other transport networks, without requiring complex IP data plane capabilities. Ongoing standardisation initiatives focus on G.8113 (Requirements for OAM functions in T-MPLS based networks) and G.8114 (Operation and maintenance mechanism for T-MPLS layer networks).

In relation to OAM, the goal of HIPT is to be able to find a clear relation between user quality of experience and quantitative measurements in order to determine the threshold for sufficient network layer QoS. The vision is then to enable the Carrier Ethernet infrastructure to detect QoS degradations below the threshold by appropriate OAM mechanisms in the network layer. The level of IPTV awareness in the network layer OAM functions is a topic of research in HIPT.

4.3 Resiliency and survivability

In PBB-TE, working and protection paths are precalculated, and the forwarding tables in nodes on these paths are provisioned/configured with the required forwarding entries. Both the working and protection traffic uses the MAC address of the destination node to fill the Destination MAC Address field. Working traffic uses the VID value assigned for the working path. A different VID value is used to forward the traffic along the protection path. Faults are detected and forwarded using a subset of 802.1ag connectivity fault management. Loss of Continuity Check is interpreted as a fault and also triggers protection switching. During path protection switching, the source nodes swap the VID value to redirect traffic onto the preconfigured protection path.

If span or local-bypass protection is used, VID tags are swapped at transit locations that bracket the failed span, node or sub-network connection. Switch-over times are short, because the required VID value and path are preconfigured.

Survivability is specific to the transport network. T-MPLS therefore defines its protection capability using ITU-T's Recommendations G.8131/Y.1382 (T-MPLS linear protection switching with 1+1, 1:1 and 1:N options) and G.8132/Y.1383 (T-MPLS ring protection switching). MPLS Fast ReRoute (FRR) capability requires the use of LSP Merge that is excluded from T-MPLS.

Network survivability traditionally deals with connection recovery after infrastructure or equipment failures, e.g. cable cuts or node outages, which are characterised by loss of signal. In addition to these "hard failures", users of IPTV may experience signal quality degradation as a "soft failure", caused by gradual component degrading or malfunction. Since the customers' perception of the signal quality is critical for the success of IP-TV, survivability measures related to "soft failures" will be investigated in HIPT.

5 Conclusion

Carrier Ethernet networks based on PBT-TE and T-MPLS are strong candidates for transport of IPTV services. The technologies offer basic OAM functionality, Managability and support for QoS required to provide Carrier grade quality. This article provided an overview of the current level of standardisation for PBT-TE and T-MPLS followed by the HIPT network architecture for transporting IPTV services over Carrier Ethernet. A number of enhancements are still required for PBT-TE and T-MPLS to make them ready for large scale IPTV deployments. The article discussed the necessary enhancements and provided possible solutions and directions that is addressed in the frame of the HIPT project.

References:
[1] IEEE Std. 802.1Q-2005: Virtual Bridged Local Area Networks.
[2] IEEE Std. 802.1ad-2005: Virtual Bridged Local Area Networks. Amendment 4: Provider Bridges.
[3] IEEE Draft 802.1ah: Provider Backbone Bridges.
[4] IEEE Std. 802.3ah-2004: Ethernet in the First Mile.
[5] IEEE Draft 802.1ag: Connectivity Fault Management.
[6] ITU-T G.805-2000: Generic functional architecture of transport networks.
[7] ITU-T G.809-2003: Functional architecture of connectionless layer networks.