A Unified Network Control and Management System for an Integrated EPON and Metro Optical Network Testbed

Yong Liu, Minh Hoang Phung, Kee Chaing Chua, Gurusamy Mohan
Yanmei Guo, Sona Panneerselvan
Department of Electrical & Computer Engineering
National University of Singapore
{eleliuy, elepmh, chuakc, elegm, elegy, elepsona}@nus.edu.sg

***Abstract*— In this paper, a unified network control and management system for an integrated Ethernet-based optical network testbed is introduced. The integrated optical network testbed consists of an Ethernet Passive Optical Network (EPON) access network and an metro Ethernet network. A unique feature of the developed network control and management system is that all network control and management signaling frames are transmitted and processed in Ethernet layer. As such, the system has a unified architecture for both the metro Ethernet network and the EPON access networks, and provide a uniform interface for network administrators to manage both networks. The architecture and the major functionalities including virtual local area network (VLAN)-based traffic engineering and resource reservation are discussed.**

I. Introduction

Ethernet protocol is ubiquitous in local area networks (LANs) due to its simplicity, high bandwidth support and the low cost of Ethernet networking equipments. Recently, it is becoming increasingly popular in access networks and metro area networks. For access networks, Ethernet Passive Optical Network (EPON) [1]–[3] provides a cost effective and high efficient solution for the last mile problem in fiber-to-the-home (FTTH) services [4]. Ethernet is also being used in metro networks [5]–[8] as a cost effective solution to support carrier grade transport services using Ethernet. Provider Backbone Transport (PBT) has been proposed by various network equipment vendors to provide connection-oriented services in Ethernet. Although there are some works on EPON testbed [1] and EPON systems are also commercially available, few testbeds consider the integration of metro Ethernet and EPON. Taking this into consideration, an integrated EPON and Ethernet-based Optical Metro Network testbed is developed in Singapore under a national Optical Network Focused Interest Group phase II (ONFIG II) program. The integrated network testbed aims to provide broadband end users with a single link carrying voice, video and services (triple-play) at lower overall communication costs. For efficient utilization and management of the integrated optical network testbed, a unified network control and management framework is also developed. This will be discussed in detail in this paper.

A unique feature of the network control and management system is that it treats the entire testbed, including the metro network and attached EPONs, as a single Ethernet. Such an approach has several advantages. Firstly, it gives end users the Ethernet interface that they are already familiar with. So one user can communicate with another across the city as if they are in the same LAN. Secondly, network administrators are presented with an uniform interface for both the metro and the EPON subsystems in the testbed. The uniform interface makes it easy to manage the entire integrated network testbed and helps save network operating cost.

The network control and management framework has the following capabilities: (i) Connection management and dynamic resource allocation to set up, manage and tear down end-to-end connections at sub-wavelength levels; (ii) Traffic Engineering under dynamic changing traffic conditions to achieve high network resource utilization; (iii) Performance management and network monitoring; and (iv) Configuration management that keeps track of network equipments in the whole network. The scope of the framework extends up to the Optical Network Unit (ONUs) of the EPONs, which are the logical separating points between end users and the network. Hence, it will be responsible for connection setup and bandwidth provisioning between a source ONU and a destination ONU. The same applies for its performance and configuration management capabilities. Anything beyond ONUs belongs to end users' domain.

The organization of the paper is as follows. In Section II, we introduce the integrated optical network testbed which includes EPON access network and metro core network. In Section III, we present the architecture for the unified network control and management system, which is followed by the description of the major functionalities in Section IV. Finally, we conclude in Section V.

II. Integrated Optical Network Testbed

A. Overview

Figure 1 shows the overall architecture of the network testbed. It consists of a metro network and a number of connecting Ethernet passive optical networks (EPONs). End users are connected to the ONUs of the EPONs. A network

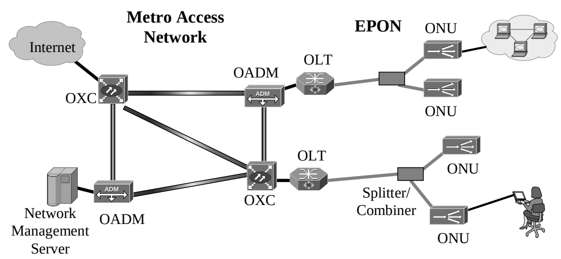

Fig. 1. Integrated optical testbed with a metro core network and EPON access networks

management server is attached to one of the core node in the metro network.

In the metro network shown in Figure 1, each node is built as an Ethernet switch on top of an optical switch which is either an Optical Add/Drop Wavelength Division Multiplexer (OADM) or an optical cross-connect (OXC). This allows data to be transported in the entire integrated network as Ethernet frames without going through the IP layer. The optical switches provide dynamic lightpath reconfiguration capability to the network. The Ethernet switches view the lightpaths as a virtual topology on which to switch data frames.

B. EPON

There are two EPON systems in this integrated testbed. Each EPON system has one Optical Line Terminal (OLT) and two ONUs. Each OLT is connected to core Ethernet switch. In each EPON system, two ONUs are connected to OLT through 1Gbps optical fibre using optical splitter/combiner. ONUs are connected to end users which will send various types of traffic.

In the EPON systems, the upstream and downstream transmission channels are separated using different wavelengths. In the downstream direction, data frames are broadcasted by the OLTs and every ONU only receives the ones destined to it. But in the upstream directions, time division media access (TDMA) is applied for the ONUs to share the upstream bandwidth. In this TDMA solution, each ONU is allocated a time slot or transmission window for data transmission by the OLT. Each time slot is capable of carrying several Ethernet frames. Packets received from one or more users are buffered in an ONU until the time slot for that ONU arrives. Upon the arrival of its time slot, the ONU will send out its buffered packets at the full transmission rate of the upstream channel during the allocated time slot. In order to avoid data frames from different ONU colliding with each other and obtain high bandwidth utilization, multi-point control protocol (MPCP) is introduced to the system. MPCP resides at the MAC control layer of the EPON system. MPCP relies on two Ethernet control messages, GATE and REPORT, to allocate bandwidth to each ONU. The GATE message is used by an OLT to allocate a transmission window to an ONU. The REPORT message is used by an ONU to report its local conditions to an OLT.

In our implementation, OLTs will send the gate message to ONUs including starting time and duration time for data transmission. An OLT grants an ONU the number of bytes it requested up to a maximum window size. This is the most conservative scheme because it assumes that no more packets will arrive after the ONU sends its request. In addition, to support QoS in EPON system, each ONU will have different queues for incoming frames with different priorities. When it is the turn for an ONU, the ONU will send first the waiting frames in the queue with higher priorities.

C. Ethernet-based Metro Network

Each node in the metro network consists of an Ethernet switch on top of an optical switch, which can either be an OADM or an OXC. The optical switches set up and release lightpaths as requested by the network management system. Electronic Ethernet frames from the Ethernet switches are converted into optical signals and transmitted on the lightpaths. At the end of the lightpaths, the optical signals are dropped, converted back to electronic and fed to the Ethernet switches.

Ethernet protocol is used for the whole metro network to simplify the interface between metro network and EPON access networks. Therefore, there is no need to translate data frames between metro network and access networks. To avoid transmission loops, spanning trees are imposed on the lightpath topology. On each spanning tree, a Virtual-bridged LAN (VLAN) is formed, which covers the entire spanning tree. All Ethernet frames are tagged by a VLAN ID before they enter the metro network. The VLAN ID will be used by Ethernet switches for loop-free data switching.

III. The Architecture of a Unified Network Control and Management System

A. Overview

Due to the Ethernet-over-optical switch architecture of the metro network and EPON-based access network, Ethernet is used as the sole transmission protocol in the entire integrated network testbed. End users will view the entire network as a unified Ethernet network. Because of this, Ethernet is also used as the sole transmission protocol for our network control and management system. In other words, all network control and management signaling frames are transmitted and processed in the Ethernet layer. This choice leads to a number of differences between our framework and traditional IP-based network control and management frameworks in approaches to traffic engineering, bandwidth reservation, etc. which will be discussed in detail in later sections.

The framework is responsible for managing network equipments in the entire testbed, including the metro network and EPONs. The management functions include keeping track of the equipments, performance management and fault management. On the control side, it provides both long-term and short-term control functionalities. Long-term control is involved with the setup and teardown of lightpaths among core nodes in the metro network. This is done based on traffic patterns and lightpath utilization statistics as reported by the nodes. When any lightpath change happens, the spanning tree/VLAN structure also has to be reconfigured. If a new

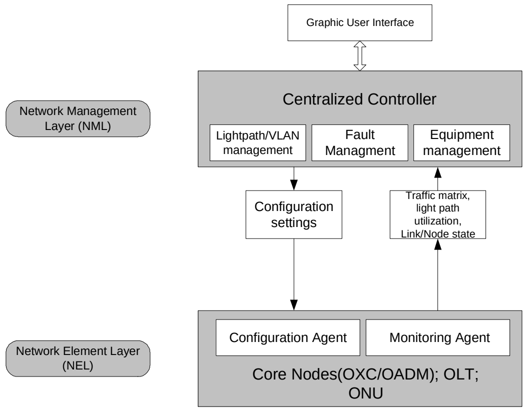

Fig. 2. Architecture of a unified network control and management system

lightpath is added, new spanning trees and VLANs have to be formed to utilize the new lightpath. If a lightpath is removed, existing traffic on the VLANs involved is migrated to other VLANs. Due to the limited number of core nodes inside the metro networks, network management and long-term control is coordinated by a centralized server. The architecture is shown in Figure 2, which consists of two logical management layers, namely network management layer (NML) and network element layer (NEL). A graphic user interface (GUI) will be present on top of the whole architecture to facilitate the network management. On the other hand, short-term control has to do with connection management such as connection setup/teardown, bandwidth reservation, etc. Since this occurs much more frequently than the above management and long-term control functions, short-term control functions are accomplished in a distributed manner among the nodes in the testbed.

B. *Network Element Layer*

The NEL consists of management agents built on top of all physical network equipments including core node (core switch and OXC/OADM), OLT and ONU. There are two types of such management agents, namely monitoring agent and configuration agent. The monitoring agent is responsible for performance monitoring, traffic measurement, link status monitoring, node or link failure detection, etc. The monitoring agent will report these information periodically to the centralized controller or upon the request of the controller. The configuration agent will conduct configuration according to the commands from the centralized controller. These configurations include lightpath establishment and elimination, etc.

C. *Network Management Layer*

The NML will be responsible the control and management of the entire network. There are three major modules for this layer which are located inside a powerful centralized controller. These modules are lightpath management module, fault management module and OLT/ONU management module. The details of these modules are described as following:

1) Lightpath/VLAN management module: This module will maintain the global and logical information on the metro networks. It will maintain the information on how the core nodes are topologically interconnected. It will also request each core node to report the relevant traffic pattern including bandwidth requirement, holding time, source and destination information, and existing lightpath utilization. Based on these information, lightpath computation algorithms inside this module will determine the optimum lightpath layout/virtual topology configuration for a given traffic pattern. This module will also consider the lightpath reconfiguration under the dynamic changing traffic conditions. The lightpath reconfiguration ability will efficiently improve the network recourse utilization. The module will also consider how to minimize the effect of lightpath reconfiguration on the exiting connections. A properly chosen trade-off will be chosen between the light-path utilization and effect on existing connections. In addition, the module will also consider the benefit to tear down an existing lightpath and set up a new one according to the changing traffic patterns reported by each core node. These lightpath configuration settings will be sent to the corresponding core nodes for light path establishment or releasing. This module will be also responsible for VLAN computation if a new lightpath is set up. The new VLAN topology will also be sent to each core node for traffic engineering purpose.

2) Fault management module: The fault management module will receive and process the link or equipment failure reported from all network elements. This module will also monitor the automatic protection switching which is implemented by each core node. This module will also be responsible for primary and backup path selection for each lightpath between any two core nodes. The latest physical network topology and virtual/lightpath topology will also be reported to the lightpath management module. All the link and equipment states will be displayed in the GUI to facilitate network management. The module will also send alarm notification to network administrator for any equipment failure events.

3) ONU and OLT management module: This module will maintain the management information on all ONUs and OLTs in the whole testbed. This module will periodically send probe packets to check the status of ONUs and OLTs. This module will be responsible for processing of the reports of lightpath utilization, traffic matrix from OLTs. This module will also provide support for traffic engineering by assigning various VLAN ID to each ONU such that each Ethernet frame will be tagged properly before they enter the integrated optical networks. This module will also send configuration commands to each OLT to configure dynamic bandwidth allocation (DBA) algorithms for the support of RSVP-based bandwidth allocation. All these functionalities will be discussed in detail in next section.

IV. Major Functionalities of the Unified System

A. *VLAN-based Control Plane*

Since all control and signaling messages are in the format of Ethernet, VLAN is used here for loop-free frame switching.

For the metro network, two dedicated VLANs will be used as the control plane for all control and management signaling messages between all core nodes and the controller. One is the working VLAN, the other one is for backup in case of link failure. These two dedicated VLANs will not bypass any core node in the whole network to enable monitoring and configuration of all core nodes. These two VLANs will be set up and selected during the network initialization stage. All control messages will be dropped and converted to electrical signals at each core node. If the current node is the destination of the signaling message, it will be processed locally. If not, it will be converted back to optical signaling and sent over the dedicated wavelength to other nodes for further processing. For the management of ONUs in the access network, all management information from each ONU will be relayed by the OLT with which the ONU is connected. All signaling messages from OLTs will also be forwarded by core nodes to simplify the network infrastructure.

B. VLAN-based Traffic Engineering

Since the data forwarding paths in Ethernet networks are based on self-learning switching tables rather than routing as in IP networks, there should be a separate route-pinning mechanism to fix the paths of end-to-end connections according to the requirements of traffic engineering algorithms. We adopt a combination of VLAN and Multiple Spanning Tree Protocol (MSTP) as defined in IEEE 802.1q-2003 standard as the route-pinning mechanism.

The approach is described as follows. Several spanning trees are first generated in the central server. The following requirements apply to the set of spanning trees generated: (i) Each spanning tree must cover all nodes of the network; (ii) Each link of the network must be included in at least one spanning tree and (iii) There must be at least two distinct paths belonging to two distinct spanning trees between every node pairs; one of them is the least-cost path. The third requirement is necessary because traffic between every node pair can be distributed to at least two different paths to achieve load balancing. After that, a VLAN is set up to cover each spanning tree. As such, each VLAN has a spanning tree topology and covers all nodes in the network.

Since EPONs already have tree topologies, all generated spanning trees will be the same in the EPON part of the integrated network. Therefore, based on MSTP, we will set up a single Common Spanning Tree (CST) for the unified network and treat the metro core network as a super node in the CST. Within the metro network, several Internal Spanning Trees (ISTs) will be generated according to the above criteria. Each VLAN topology will now consist of an IST inside the CST.

The ISTs are built as follows. The number of ISTs is equal to the number of nodes in the core networks. Each core node is the root of one IST. This method has two advantages. Firstly, it balances traffic loads in the core network. This is because traffic within each spanning tree tends to concentrate at the root due to the spanning tree topology. Secondly, there exists at least one least-cost path between every core node pair in the two spanning trees whose roots are those two nodes.

C. RSVP-based Bandwidth Allocation and QoS Guarantee

Since it is impractical and unscalable to require each user application to make explicit bandwidth reservation, a bandwidth reservation module is developed to perform the task. The bandwidth reservation module is running at each ONU, which makes bandwidth reservation on behalf of end users. In this case, the reservation module continuously monitors traffic passing through the ONU and detects the presence of real-time traffic, which is recognized by the protocol field in the IP header. When real-time traffic is detected, the module measures the data rate and reserves enough bandwidth between the local ONU and the destination ONU to provide QoS guarantee for real-time traffic.

Due to the way agents are connected in the control plane, an RSVP-like protocol is used for bandwidth reservation. In this protocol, the source sends a Resv message that contains the requested bandwidth, the VLAN ID to be used, and the MAC address of the destination to the server. In the specified VLAN, the server does a simulated routing and determines the intermediate nodes on the path between the source and the destination. It then relays the Resv message to those nodes. The nodes check if enough bandwidth is available to satisfy the request and report back to the server. If all the intermediate nodes report success, a success message is sent to the source and destination. Otherwise, an error message is sent. During transmission, the source is required to refresh the reservation periodically using the same protocol. The refresh Resv message contains a flag for nodes to distinguish it from new reservation.

D. VLAN Tag Encapsulation

Based on the network topology of the testbed, the server determines a number of VLANs and remotely configures those VLANs onto the core switches through configuration agents in the core switches. The VLANs generated should satisfy the requirements in Section IV-B. Once the VLANs are generated, the server determines a set of VLAN IDs (comprising at least two VLANs) for each pair of core Ethernet switches to communicate with each other. This VLAN ID set is sent to all ONUs in the EPONs connected to Ethernet switches. Each ONU randomly allocates each user connected to it a VLAN ID. This allocated VLAN ID will be used for all future communication with users connected to the other Ethernet switch. All users data will be tagged by the assigned VLAN ID at the ONU before enter the integrated optical network.

Two encapsulation schemes [11] can be used for each frame entering the unified network which are shown in Figure 3.

- MAC address encapsulation (M-in-M): The goal of the unified network is to act as a virtual Ethernet LAN to the end users. However, the number of end users is likely to be much larger than the capacity of switching tables of the nodes. To prevent this MAC address explosion, upon receiving a frame from end users, the source ONU inserts

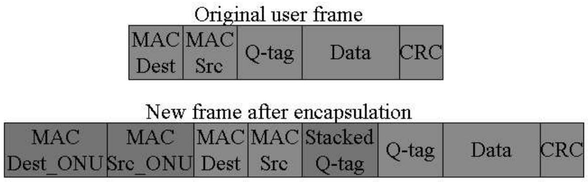

Fig. 3. Frame encapsulation schemes

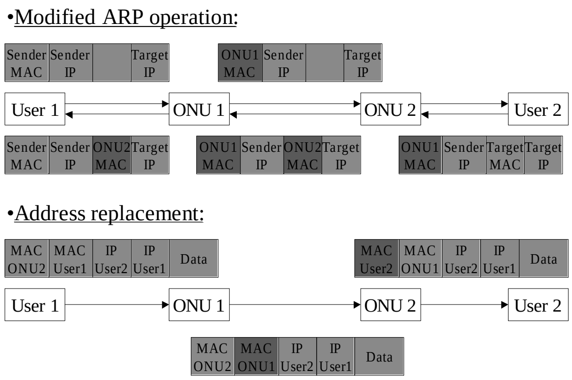

Fig. 4. MAC address replacement scheme

- its address and destination ONU address into the frame for switching purposes in the network. These addresses are removed by the destination ONU.
- VLAN ID encapsulation (Q-in-Q): VLAN may be utilized by local network administrators for various purposes. In this approach, the VLAN ID assigned to each connection is stacked on top of those used by local networks. This is done at the core node closest to the source node. The stacked VLAN ID is removed at the destination core node.

The above encapsulation schemes may make the frame size exceed the MTU. Therefore, a new address replacement scheme is developed for our testbed. In this scheme, the source and destination MAC addresses are replaced by ONU MAC addresses through modification of address resolution protocol (ARP) before the frame is sent into the metro network. The operations involved are described in Figure 4.

In our modified ARP process, User 1's MAC in the probe frame is replaced by ONU1's MAC at ONU 1. Similarly, User 2's MAC in the acknowledgement frame is replaced by ONU 2's MAC at ONU2. Therefore, User 1's IP and user 2's IP are associated with ONU 1's MAC and ONU 2's MAC, respectively. Consequently, data frames destined for User 2 are switched to ONU 2 by the integrated optical network. ONU 2 will maintain a look up table to store MAC addresses of all connected end users. After data frames destined for User 2 arrives, ONU 1 will replace ONU 2's MAC with User 2's MAC and forwards the data frames to User 2.

As a result of this scheme, User 1 and User 2 are represented by ONU 1 and ONU 2's MAC addresses in the integrated network testbed.

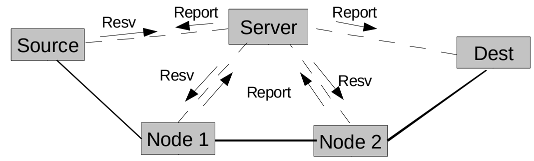

Fig. 5. The architecture of a unified network control and management system

Stacked VLAN tag is eliminated by reserving a portion of VLAN ID space for network use. Due to the large ID space (4096 possible IDs), this will not have impact on user's VLAN usage.

E. Implementation Details

1) Hardware limitations: For the integrated optical testbed, we have following main limitations on hardware devices.

- Ethernet switches: There is no support for automatic spanning tree configuration (either STP or MSTP). In addition, communication with software agents in a switch is only possible through a separate management port at the back of the switch, but not through normal data ports.
- EPON: There is no access to the internal working of either OLTs or ONUs. The only QoS service offered is strict priority differentiation. In this QoS model, frames marked as high priority will be forwarded before low priority ones by the ONUs.

These limitations necessitate several changes in implementation plan. They are described in the following section.

2) Implementation details: Two main tasks required from EPON in our framework are frame header modification and bandwidth reservation. Due to limitation of EPON, these tasks are done by some control PCs attached to the EPON, namely ONU_PCs and an OLT_PC. An ONU_PC is inserted in the data plane between every ONU and its end users. ONU_PCs perform frame header modification as described earlier. In addition, they perform bandwidth reservation together with OLT_PC. An OLT_PC is inserted between every OLT and the connected Ethernet switch.

Due to limitation on Ethernet switches, a separate control plane is introduced to carry signaling information between the server and software agents. This control plane is composed of Ethernet cables connecting the management ports of Ethernet switches, management interfaces of ONU_PCs, the central server and OLT_PCs to a central hub. As such, the control plane comprises a single broadcast domain through which any network entity can communicate directly with any other entity.

VLANs are calculated by the central server according to the virtual topology provided by lightpaths instead of physical links. In this method, the central server must have knowledge of the virtual topology. The server calculates spanning trees based on this topology information and allocates one VLAN to each spanning tree. VLAN configuration commands that include VLAN ID, port information are then sent to the configuration agents running on Ethernet switches.

Due to the separate control plane, bandwidth reservation is carried out as shown in Figure 5. Reservation requests are sent to central server through the control plane. The server determines which nodes are on the path between the source and destination and relays the requests to them. The nodes check the available bandwidth against the requested amount. If enough bandwidth is available, the request is granted. Otherwise, it is denied. The result is reported back to the server, which relays it to the ONU_PCs at the source and destination. The ONU_PC is responsible for marking bandwidth-guaranteed frames as high priority before they enter the ONU. The ONU_PC is also responsible for policing its traffic, i.e. ensuring that the rate of outgoing high priority traffic does not exceed the reserved bandwidth.

V. Conclusions

In this paper, we have introduced a unified network control and management system for an integrated EPON and Metro optical network testbed. We introduce the integrated optical network testbed which includes EPON access network and metro core network. Then, we describe the architecture for the unified Ethernet-based network control and management system. The major functionalities including VLAN-based control plane, VLAN-based traffic engineering, RSVP-based bandwidth allocation, and VLAN encapsulation are also discussed in details, which can be used a useful reference for further development of integrated ethernet-based optical network control and management systems.

References

[1] T.P. Lin and K.P. Fan, "EPON testbed and field trial environment in Taiwan," TRIDENTCOM 2006.

[2] G. Kramer, B. Mukherjee and G. Pessavento, "Ethernet PON (ePON): design and analysis of an optical access network," Photonic Network Communications, Vol.3, No.3, pp.307-319, 2001.

[3] M.P. McGarry, M. Maier and M. Reisslein, "Ethernet PONs: a survey of dynamic bandwidth allocation (DBA) algorithms," IEEE Communications Magazine, Vol.42, No.8, pp. S8-15, 2004.

[4] K. Kim, "The evolution of PON-based FTTH solutions," Information Sciences, Vol.149, No.1-2, pp.21-30, 2003.

[5] M. Ali, G. Chiruvolu, A. Ge, "Traffic engineering in metro Ethernet," IEEE Network, Vol.19, No.2, pp.10-17, 2005.

[6] X. He, M. Zhu, Q. Chu, "Traffic Engineering for Metro Ethernet Based on Multiple Spanning Trees," International Conference on Systems and Mobile Communications and Learning Technologies 2006.

[7] http://www.metroethernetforum.org

[8] http://www.netadmin.se

[9] Y.M. Yang, J.M. Nho, N.P. Mahalik, K. Kim and B.H. Ahn, "QoS provisioning in the EPON systems with traffic class burst polling based delta DBA," IEICE Transactions on Communications, E89-B(2):419-426, 2006.

[10] S.R. Sherif, A. Hadjiantonis, G. Ellinas, C. Assi and M.A. Ali, "A novel decentralized ethernet-based PON access architecture for provisioning differentiated QoS," Journal of Lightwave Technology, Vol.22, No.11, pp.2483- 2497, 2004.

[11] G. Chiruvolu, A. Ge, D. Elie-Dit-Cosaque, M. Ali and J. Rouyer "Issues and approaches on extending ethernet beyond LANs," IEEE Communications Magazine 2004.