國立交通大學

電機學院　電機與控制學程

碩士論文

應用於ATCA接取設備之電信等級IGMP代理伺服器

Carrier-Grade IGMP Proxy Server for ATCA-based Access Equipment

研究生：徐彬海

指導教授：蘇朝琴 教授、廖維國 教授

中華民國九十七年三月

應用於ATCA接取設備之電信等級IGMP代理伺服器

Carrier-Grade IGMP Proxy Server for ATCA-based Access Equipment

研究生：徐彬海　　Student : Ping-Hai Hsu

指導教授：蘇朝琴 教授、廖維國 教授

Advisors : Chau-Chin Su, Wei-Kuo Liao

國立交通大學

電機學院 電機與控制學程

碩士論文

A Thesis

Submitted to College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Electrical and Control Engineering

March 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年三月

應用於 ATCA 接取設備之電信等級 IGMP 代理伺服器

學生：徐彬海

指導教授：蘇朝琴 教授、廖維國 教授

國立交通大學 電機學院 電機與控制學程碩士班

摘要

如何提高系統與服務的可用度，以達到電信等級的水準，對電信設備供應商與服務供應商來說都是一個令人困擾的難題。本論文採用先進通訊電腦架構(ATCA)的硬體平台，並以乙太被動光纖網路(EPON)的接取設備為例，探討在 IPTV 的服務架構中，如何提升 ATCA-based 局用接取設備上 IGMP 代理伺服器的可用度。目前常用的乙太網路恢復技術，如：Rapid Spanning Tree Protocol (RSTP)及 Resilient Packet Ring (RPR)等，仍有過於複雜與拓墣結構的限制等問題，並不適用於 ATCA-based 接取設備。

本論文提出以負載分配、快速失效備援與無縫復原的機制，來提升 ATCA-based IGMP 代理伺服器的可用度。尤其是針對 ATCA 結構及 multicast 封包的獨特性，提出較簡易又快速的方法來達到備援同步與封包導向的目的，這些機制都大幅增進系統失效備援的效能。在本論文最後，也提出 IGMP 代理伺服器失效備援的速度的量測原理與方法，其模擬結果平均約 45ms，符合電信等級要求的 50ms。

關鍵字: 電信等級、先進通訊電腦架構、失效備援

Carrier-Grade IGMP Proxy Server for ATCA-based Access Equipment

Student: Ping-Hai Hsu

Advisors: Dr. Chau-Chin Su, Dr. Wei-Kuo Liao

Degree Program of Electrical and Computer Engineering

National Chiao Tung University

Abstract

How to increase system or service availability and to reach carrier grade is always a grand challenge for telecom equipment venders and service providers. The thesis adopts Advanced Telecom Computing Architecture (ATCA) as a hardware platform and Ethernet Passive Optical Network (EPON) as an example to research how to improve the availability of IGMP proxy server on an ATCA-based central office access equipment for IPTV application. There are lots of Ethernet resiliency technologies had been developed, such as Rapid Spanning Tree Protocol (RSTP) and Resilient Packet Ring (RPR) etc. They still have an over complication and topology limitation problems for an ATCA-based access equipment.

The thesis proposes using load sharing, fast failover, and seamless resuming mechanisms to improve the availability of IGMP proxy server. The proposed methods take the characteristics of ATCA-based topology and multicast traffic to simplify and to speed up the standby synchronization and traffic redirection. These improvements immediately shorten the failover time of IGMP Proxy and improve system availability. Finally, the thesis also proposes the failover time measurement principle and method. The average simulation result is around 45ms. It is better than 50ms, which is the carrier-grade failover requirement specified by SONET/SDH.

Keyword: carrier-grade, ATCA, failover

誌謝

一邊工作、一邊讀書真是一件不容易的事，尤其在我離開學校十年後，更顯得艱辛。

首先，我最想感謝的是我的妻子懿秀，感謝她在我求學的這段期間，獨自負責料理家中的大小事務，感謝她的付出與忍耐；另一個過意不去的就是可愛的小女安琪了，爸爸常常因為工作與學業兩邊的壓力，沒辦法陪妳玩，讓你養成看電視的壞習慣。

在此特別感謝我的指導教授 蘇朝琴 與 廖維國 老師，感謝老師們不厭其煩地指導。也感謝鴻文學長，幫我解答疑難雜症，協助我找尋研究的方向。

最後，還要感謝一些我工研院的工作夥伴：恕康與港喬，有你們的協助，一切都變得順利許多。另外還有長官寶哥，感謝你的鼓勵與支持。

沒有這些支持我的家人、學校的師長與工研院的夥伴們，是不可能有今天的我，再次獻上我最誠摯的謝意。

徐彬海
2008/03/11

Table of Contents

Lists of Figures

Lists of Tables

Chapter 1

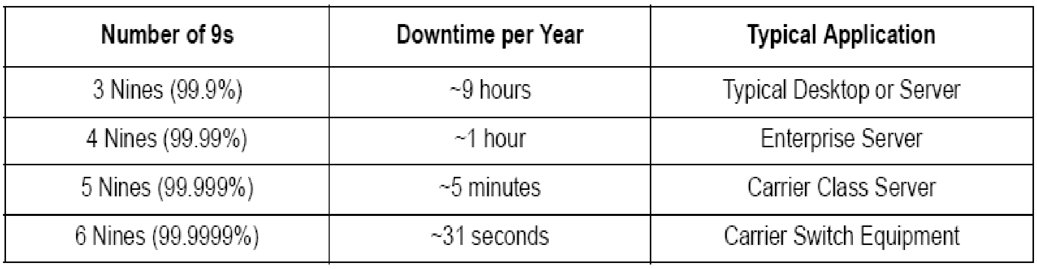
Introduction

In this chapter, we introduce the concepts and definitions of carrier grade and availability, and brief the general approaches for availability improvement. We also point the problem out about IPTV service, and to explain the motivation, goal, and solutions for our research. The last paragraph shows the organization of this thesis.

1.1 Introduction

Recently, more and more high-bandwidth services such as high-speed Internet access, IPTV, and HDTV and so on are emerging in life of people. IPTV service is surely the most important of killer application in the coming tens of years. To provide a high-availability and carrier-grade end-to-end IPTV service is always the goal for all IPTV service providers [1].

The term of "high availability" is frequently used when referring to a system that is capable of providing service most of the time. This is typically quantified in terms of the number of "9s". Table 1-1 shows the annual downtime and typical applications for various classes of systems. A system is classified as "carrier-class" or "carrier-grade" should exceed "5 nines" availability performance.



Number of 9s	Downtime per Year	Typical Application
3 Nines (99.9%)	~9 hours	Typical Desktop or Server
4 Nines (99.99%)	~1 hour	Enterprise Server
5 Nines (99.999%)	~5 minutes	Carrier Class Server
6 Nines (99.9999%)	~31 seconds	Carrier Switch Equipment

Table 1-1 Classes of high availability systems

According to the definition by Service Availability Forum (SAF), the availability is expressed by the following formula [2]:

$$Availability = \frac{MTTF}{MTTF + MTTR} \qquad (1\text{-}1)$$

Where the mean time to failure (MTTF) is the interval in which the system or element can provide service without failure. It is represented as a reciprocal of the statistical mean elapsed time to its projected or observed failure. Another attribute related to reliability is the mean time to repair (MTTR). This attribute represents the interval in time it takes to resume service after a failure has been experienced.

In general, there are three approaches to improve system availability. The first one is parallel operation. In Fig. 1-1, it shows two components, part X, operating in parallel if the combination is considered failed when both parts fail. The combined system is operational if either is available. The combined availability is shown by the 1-2. Obviously, the parallel operation is an efficient way making a highly reliable system from low reliability. i.e., all mission critical systems are designed with redundant components.

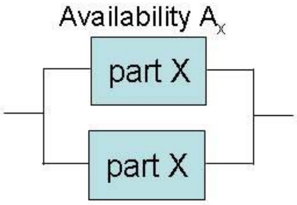


Fig. 1-1 Parallel operation configuration

$$A = 1-(1-A_x)^2 \tag{1-2}$$

The second and third ones are reducing MTTR and increasing MTTF. They are also very quite obvious form 1-1. However, to design a system with very high MTTF is a very expensive approach. It is often limited to special industries and applications, such as avionics, life-support, military, and aerospace programs. In this research, we adopt parallel operation and reducing MTTR approaches to improve system availability.

In Fig. 1-2, it is the network architecture of IPTV. It can be simply divided into two sections by which multicast protocol is adopted between customers and headhends or service providers. One is a multicast routing protocol section, and the other is Internet Group Management Protocol (IGMP).

In order to improve the availability of end-to end IPTV service, these two sections should support some self-resilient capabilities by themselves. There are already lots of researches about the fault-tolerant Multicast Routing algorithms [3-10] are developed to improve both bandwidth efficient and robust transport. Unfortunately, there are few researches to address the fault-tolerant problem for the IGMP part.

To go deep into the IGMP section, it is mainly composed of access equipments, such as Optical Line Terminal (OLT) system at central office side and Optical Network Unit (ONU) at customer side for FTTx (Fiber To The x) application. Generally, IGMP proxy function is suitable implemented on OLT system and IGMP snooping on ONU [11].

In our research, we will adopt Advanced Telecom Computing Architecture (ATCA) as our hardware platform, and Ethernet Passive Optical Network (EPON) OLT system as our application. In an ATCA-based EPON OLT system as shown in Fig. 1-3, IGMP proxy function is mainly realized by the cooperation of IGMP proxy protocol handler and Ethernet switching fabric on switch blade.

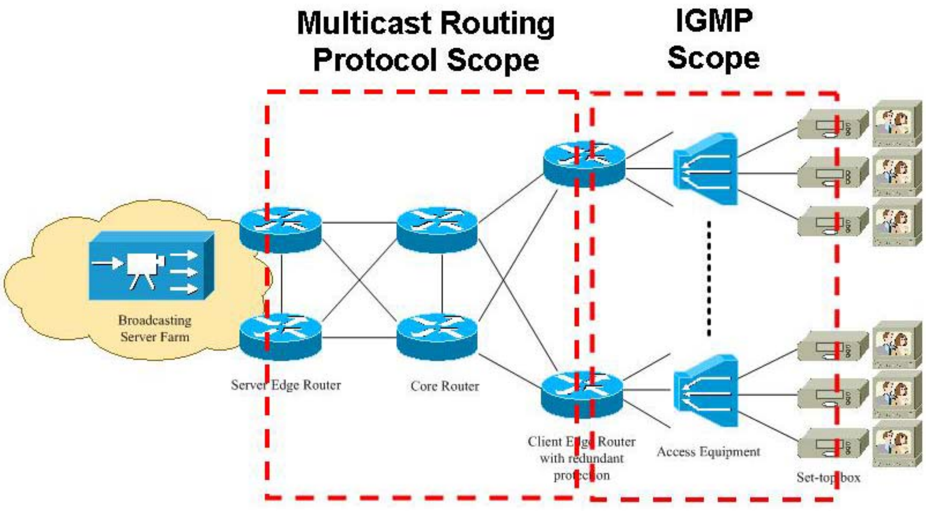


Fig. 1-2 The network architecture of IPTV

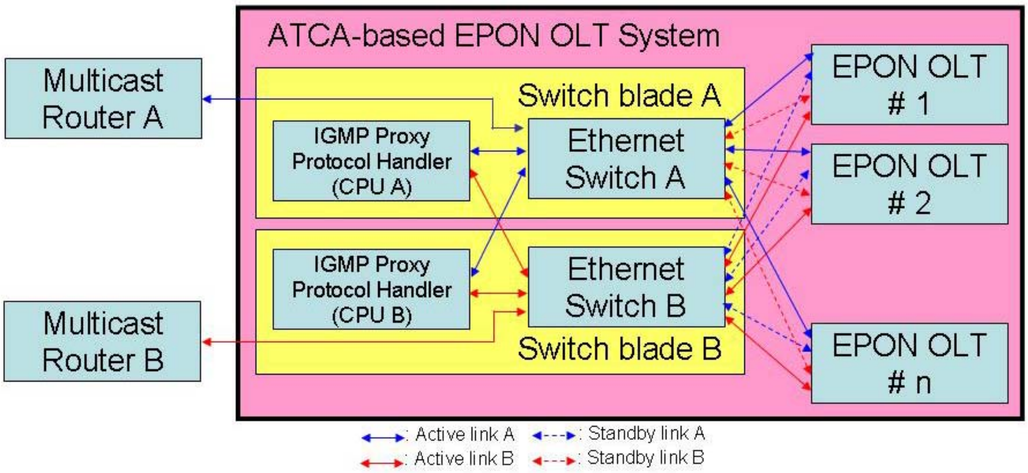


Fig. 1-3 The architecture of ATCA-based EPON OLT system

The objective of this research is to improve the availability of IGMP proxy function on an access equipment by using parallel operation and reducing MTTR approaches. The concrete method is just to develop a redundant protection mechanism for IGMP proxy with load sharing, fast failover, and seamless resuming capabilities. The Figs. 1-4 and 1-5 depict the IGMP Proxy function operational concept for normal and failover cases, respectively.

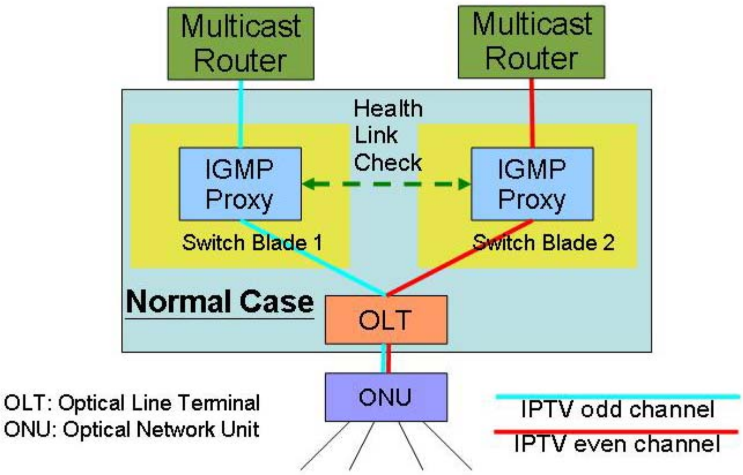


Fig. 1-4 IGMP proxy function operational concept in normal case

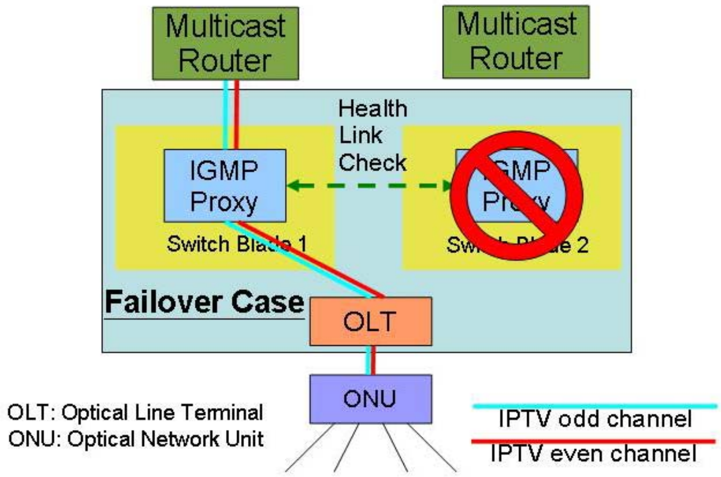


Fig. 1-5 IGMP proxy function operational concept in failover case

1.2 Thesis Organization

This thesis comprises five chapters. This first chapter illustrates the research motivation and objective. The Chapter 2 describes the background and related works. We introduce the fundamental concept about AdvancedTCA, which is the hardware platform the research adopts, IGMP and IGMP proxy, and Customer-VLAN and Service-VLAN. In addition, we compare five Ethernet resilience mechanisms in this chapter.

In Chapter 3, we describe my proposed method, which covers from design requirements, system architecture, operational scenarios, VLAN design, standby synchronization, and fast redirection.

In Chapter 4, we describe the key performance index (failover time) of this research and depict measurement principle, testing environment, and the results.

Finally, Chapter 5 concludes this thesis and discusses the future development.

Chapter 2

Background and Related Works

In this chapter, we introduce the following background technologies for our research.

- Advanced Telecommunications Computing Architecture (ATCA) platform
- Ethernet resilience mechanism survey
- IGMP, IGMP proxy, and IGMP snooping functions
- Customer-VLAN and Service-VLAN concept

2.1 AdvancedTCA (ATCA)

Advanced Telecommunications Computing Architecture, known as ATCA, is a new system form factor defined by the PCI Industrial Computers Manufacturers Group (PICMG) to provide an industry standard platform that enables building telecommunication grade products in a multi-vendor compatible environment. ATCA is the first standardized platform for high availability system with redundant power, cooling, and high-speed interconnections for data and control plane.

The ATCA specification defines two separate interconnect fabrics: the base interface and fabric interface. The base interface accommodates essential interoperability over a switched fabric supporting 10/100/1000-Mbit/s Ethernet in a dual-star configuration. The fabric interface, on the other hand, allows for full mesh interconnect architectures, high-speed switched fabric architectures, or combinations of both [12].

Fig. 2-1 illustrates an ATCA backplane implementing the base and fabric interfaces. Designed for an EIA 19-in. rack, this backplane can support up to 14 ATCA-compliant slots, 2 Switch blades & 12 Applications line cards.

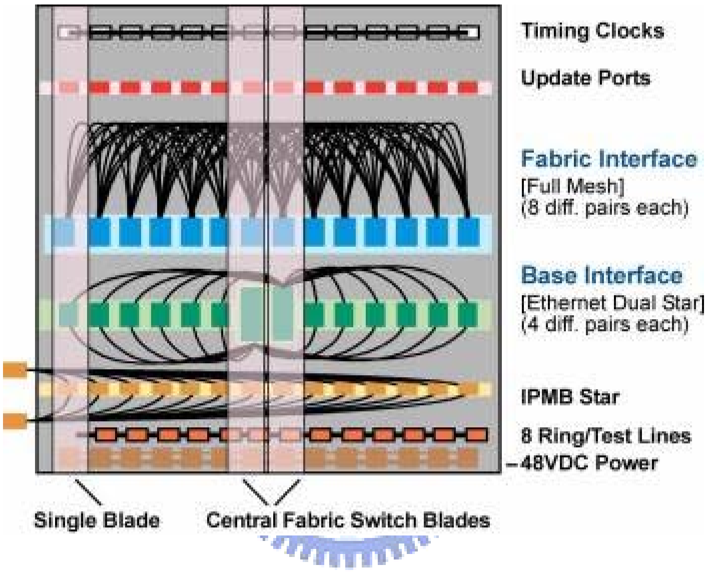


Fig. 2-1 ATCA backplane

The design choice among available topologies and interconnect protocols is governed by the targeted applications and their bandwidth requirements, by industry acceptance of a given interconnect technology, and by overall cost considerations.

In this thesis, the Fabric interface of the platform adopts dual-star topology and multiple Gigabit Ethernet interconnection technology. The platform architecture is like a 2-stage Ethernet switches one as Fig. 2-2 shown.

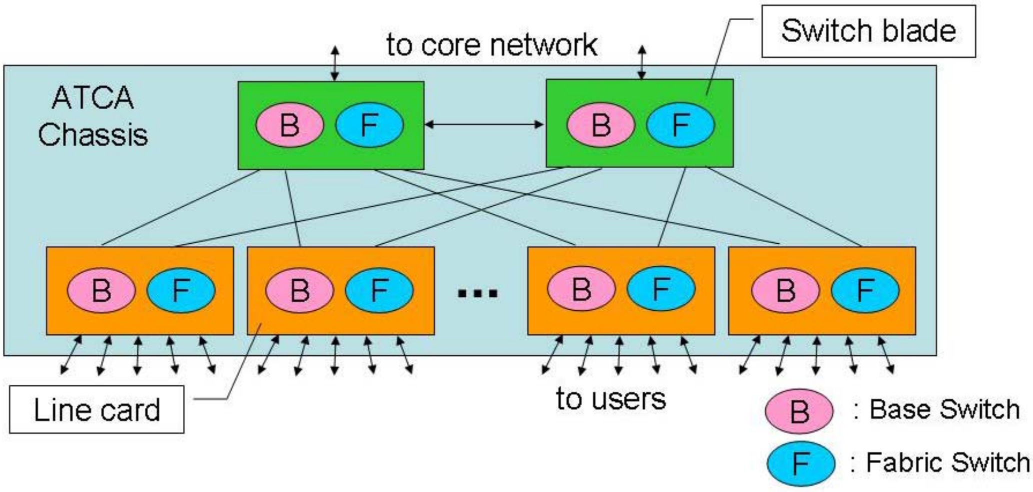


Fig. 2-2 2-stage Ethernet switch architecture

2.2 Ethernet Resilience Mechanisms

Many Ethernet technologies to improve Ethernet reliability have been proposed and some are now in use. In the standardized technologies of IEEE 802, Spanning Tree Protocol (STP) [13], Rapid STP (RSTP) [14] , and Resilient Packet Ring (RPR) [15] are specified in IEEE 802.1D, 1w, and 802.17. Vendors have also proposed proprietary technologies. Extreme Standby Routing Protocol [16] is a node redundancy technology for a tree topology network. For ring networks, there is Extreme Automatic Protection Switching [17].

STP was originally developed to solve the fundamental problem of traffic loops, i.e. it avoids logical loops in a network by blocking the traffic on some links. STP was specified in IEEE 802.1D editions prior to 2004. STP is based on time-outs to discover failures in a network, i.e. to check that special test packets arrive within a certain amount of time. These time-outs are fairly large and in practice the recovery time using STP is around 30 seconds.

RSTP is an evolution of STP to increase the performance of network recovery.

RSTP recovery is no longer based on the time-outs found in STP but includes a set of new features to achieve lower recovery times. In RSTP a failure/recovery detection mechanism is utilized, a handshaking mechanism to rapidly agree on new ports states has been included and the edge port concept for the simple case with nodes at the end of the network has been introduced. These mechanisms result in the recovery speed is very much faster than in STP.

Resilient Packet Ring (RPR), also known as IEEE 802.17, is a standard designed for the optimized transport of data traffic over fiber rings. Its design is to provide the resilience found in SONET/SDH networks (50ms protection) but instead of setting up circuit oriented connections, providing a packet based transmission. This is to increase the efficiency of Ethernet and IP services.

RPR works on a concept of dual counter rotating rings called ringlets. These ringlets are set up by creating RPR stations at nodes where traffic is supposed to drop, per flow (a flow is the ingress and egress of data traffic). RPR uses MAC (Media Access Control protocol) messages to direct the traffic, which traverses both directions around the ringlet. The nodes also negotiate for bandwidth among themselves using fairness algorithms, avoiding congestion and failed spans. The avoidance of failed spans is accomplished by using one of two techniques known as “steering” and “wrapping”. Under steering if a node or span is broken all nodes are notified of a topology change and they reroute their traffic. In wrapping the traffic is looped back at the last node prior to the break and routed to the destination station.

RPR convergence time consists in the change notification which is around one round trip time plus processing time, and the topology stabilization time (40 milliseconds by default).

Table. 2-1 compares the restoration time, topology, Standards, and layer of resiliency among these five Ethernet Resilience technologies.

Technology	Restoration Time	Topology	Standards	Layer of resiliency
STP	30 ~ 50 sec [18]	mesh	IEEE 802.1D 1998	Layer 2
RSTP	200ms ~ 2 sec [19]	mesh	IEEE 802.1w IEEE 802.1D 2004	Layer 2
RPR	40 ms (default) [18]	Ring	IEEE 802.17	Layer 2
EAPS	< 50ms [20]	Ring	RFC 3619	Layer 2
ESRP	2 ~ 6 sec [21]	Dual-star	Extreme Network (proprietary)	Layer 2/3

Table 2-1 Comparison for Ethernet resilience mechanisms

2.3 IGMP & IGMP Proxy Server

2.3.1 IGMP Overview

The section below gives a brief overview about IGMP fundamental concept when used in IPTV architecture. Then, introduces the IGMP three different versions IGMPv1 (RFC 1112) [22], IGMPv2 (RFC 2236) [23] and IGMPv3 (RFC 3376) [24].

Basic IGMP operation involves two devices:

IGMP host (or client), which issues messages to join or leave a multicast group. The client also responds to queries from the multicast router. A set-top box is an example of an IGMP host.

- IGMP router (or multicast router), which responds to the join and leave messages to determine if multicast groups should be forwarded out an interface. Periodic queries are used to recover from error conditions and verify requests. The IGMP router receives multicast groups either through the use of a multicast protocol such as PIM or via static flooding. It is the termination point for IGMP messages, so does not send any IGMP information to its upstream neighbors.

For this discussion, think of the STB as the IGMP host and the BSR as the IGMP router. IGMP provides four basic functions for IP multicast networks:

- JOIN: An IGMP host indicates that it wants to receive information from （“become a member of”）a multicast group.
- LEAVE: An IGMP host indicates that it no longer wishes to receive information from a multicast group.
- QUERY: An IGMP router can ask the hosts which groups they are members of. This is done to verify a JOIN/LEAVE request or to look for error conditions. For example, a set-top box may be been unplugged so did not issue a LEAVE command. Queries may be:
    - Specific Query: Asks whether the host is a member of a specified multicast group.
    - General Query: Asks the host to indicate all groups that it belongs to
- MEMBERSHIP REPORT: An IGMP host tells the IGMP host what groups it belongs to. This report can be either:
    - Solicited Membership Report: Sent in response to a QUERY
    - Unsolicited Membership Report: Initiated by the client

In an IPTV network, each broadcast television channel is an IP multicast group. The subscriber changes the channel by LEAVE-ing one group and JOIN-ing a different group.

The following paragraph introduces the major characteristics, enhancements and differences among the three different IGMP versions.

IGMP version 1 (IGMPv1 - RFC 1112) is not used for IPTV, because it does not include an explicit “LEAVE” capability. The client will continue to receive all requested streams until the multicast router query timeout.

IGMP version 2 (IGMPv2 - RFC 2236) and version 3 (IGMPv3 - RFC 3376) can both be used for IPTV. Like its predecessor, IGMPv2 supports Any Source Multicast (ASM) networks. In an ASM network, the IGMP host specifies the multicast group that it wishes to join, and receives all traffic with the specified multicast address regardless of who is sending the traffic. Most deployed IPTV clients (set top boxes) support IGMPv2.

The major enhancement in IGMPv3 is support for Source Specific Multicast (SSM). When using SSM, the host specifies the source address that it will listen to. This is an important security enhancement since it prevents clients such as a set top box from receiving traffic generated by other subscribers on the network. IGMPv3 is backwards-compatible with IGMPv2.

2.3.2 IGMP Proxy Server

IGMP Proxy Server is especially suitable on an edge aggregation box such as a Digital Subscriber Line Access Multiplexer (DSLAM) with a tree topology which an edge box has only one connection to the core network side and has many connections to the customer side. The behavior of an IGMP Proxy Server is standardized in RFC 4605. It is described as below.

A device performing IGMP proxy server acts in a dual mode as an IGMP router and IGMP client as shown in Fig. 2-3.

When the IGMP host issues a join message, the proxy server will receive the join and add the interface to its outgoing interface list for a specific multicast group. A General Membership Query timer and state will be used by the proxy server to send general queries downstream to all multicast enabled interfaces. When a leave is received, the proxy server will be responsible for issuing a group-specific query and removing the interface from the outgoing interface list if no hosts respond within the configured response time interval. When interacting with the IGMP hosts, the proxy server appears as an IGMP router.

If the proxy server is already receiving a group from an upstream router it will not issue a join message upstream. However, if a downstream host joins a group not currently received by the proxy server, the proxy server will issue its own join upstream to the multicast router. When a group is no longer required by downstream hosts, the proxy server will issue a leave message upstream to stop the flow of packets destination for that multicast group. The proxy server will also respond to Membership Queries sent from the multicast router. When interacting with the multicast router, the proxy server appears as an IGMP client.

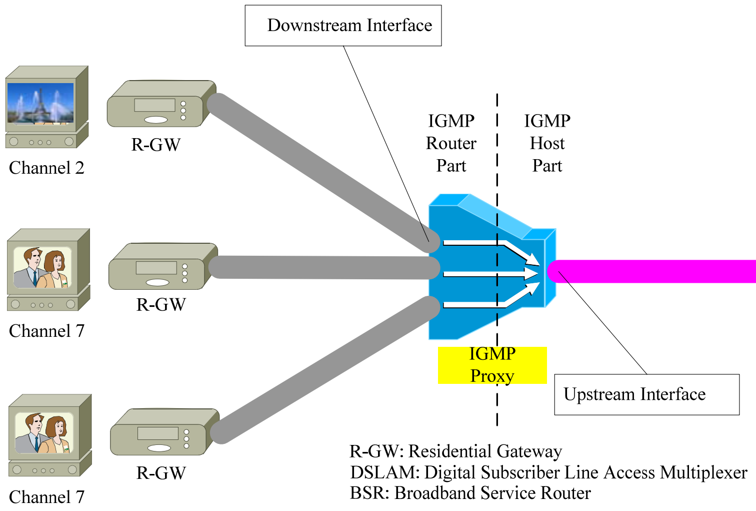


Fig. 2-3 IGMP proxy server

2.4 Customer VLAN and Service VLAN

There are two fundamental VLAN design options:

Customer VLAN: In this model, there is a dedicated VLAN for each subscriber. This is also called the 1:1 model since there is one VLAN per subscriber.

Service VLAN: In this model, there is a dedicated VLAN for each service. This is also called the N:1 model since multiple subscribers share each VLAN.

2.4.1 Customer VLAN

In the Customer VLAN (C-VLAN) model, there is a dedicated VLAN for each subscriber. The C-VLAN is created between the access elements (e.g. ONT/OLT, DSLAM) and the BSR, and carries all traffic for all services to and from an individual subscriber.

Fig. 2-4 shows a typical C-VLAN scenario. The 1:1 mapping of VLANs to customers enables the edge router to effectively manage the bandwidth for each subscriber.

The drawback to the 1:1 mapping is that C-VLANs are point-to-point paths which do not natively support distributed multicast replication between C-VLANs. As illustrated in Fig. 2-4, the edge router would need to create a unique stream for each user, even those watching the same channel. This requires additional bandwidth between the DSLAM and the edge router. The DSLAM cannot replicate channels since it does not realize that the same channel is coming in over multiple virtual connections.

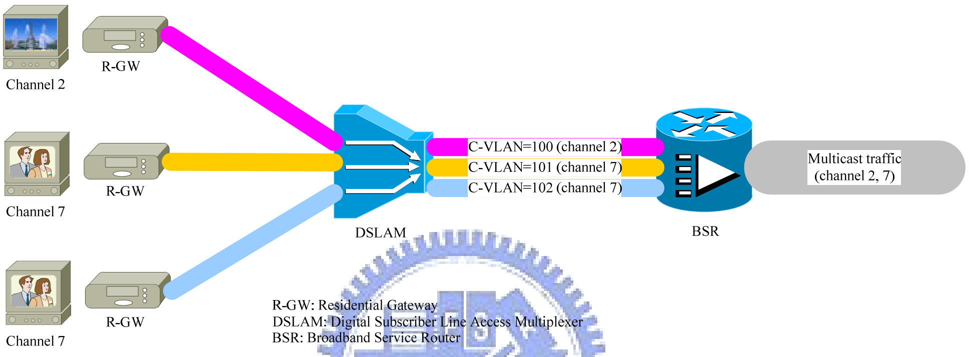


Fig. 2-4 Multicasting with customer VLAN

2.4.2 Service VLAN

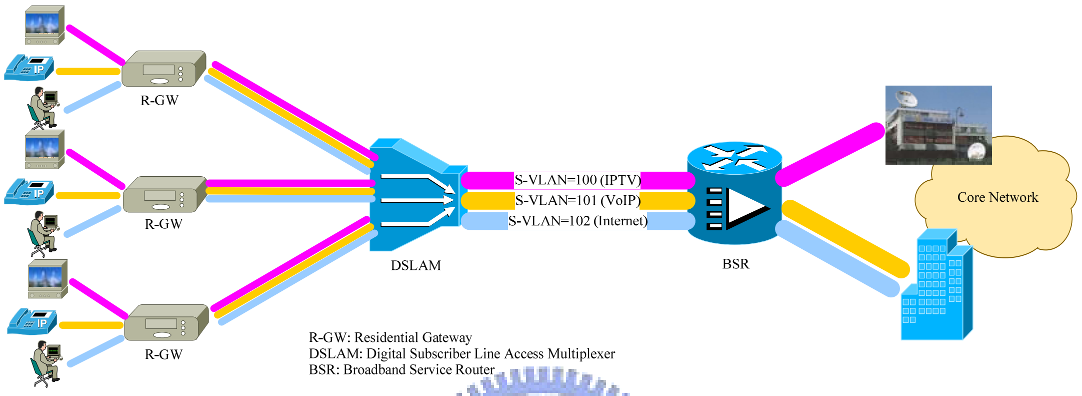
In the Service VLAN model, there is a shared VLAN used to deliver services to subscribers. A separate VLAN is used for each service. This is depicted in Figs. 2-5 and 2-6.

This architecture is frequently used when IPTV service is introduced onto an existing broadband network. Putting new services into a different VLAN lowers the risk of disrupting the existing service. It is especially useful in two situations:

Wholesale networks, where a third party (“wholesaler” or “virtual network operator”) provides some of the services (such as IPTV). In this case, the VLAN carrying the wholesaled service can be delivered across the backbone to the wholesaler is data center as shown in Fig. 2-5.

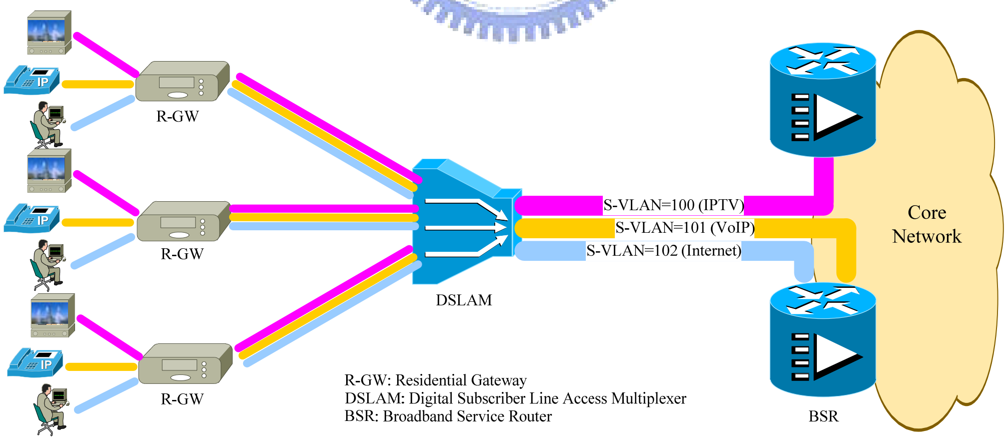
Multi-edge networks, where a separate edge router is installed to handle a new service. In this case, the VLAN for the new service can terminate at a separate edge router. This is depicted in Fig. 2-6.

.



R-GW: Residential Gateway
DSLAM: Digital Subscriber Line Access Multiplexer
BSR: Broadband Service Router

Fig. 2-5 Service VLAN model (Wholesaler)



R-GW: Residential Gateway
DSLAM: Digital Subscriber Line Access Multiplexer
BSR: Broadband Service Router

Fig. 2-6 Service VLANs model (Multi-Edge)

2.4.3 Hybrid Model

The hybrid model leverages the strengths of both architectures, creating a single policy enforcement point while providing efficient multicast delivery. The hybrid model leverages multiple VLANs as follows:

A subscriber-dedicated C-VLAN carries unicast traffic such as Internet Access, Voice over IP and Video on Demand between the access node and the BSR.

A shared Service VLAN carries broadcast television traffic to each DSLAM. In this special case, this is called the Multicast VLAN (MC-VLAN).

This is depicted in Fig. 2-7.

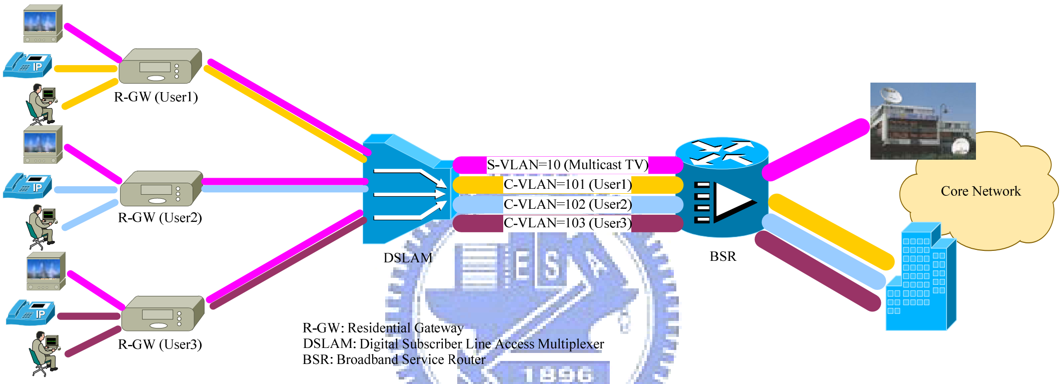


Fig. 2-7 Hybrid VLAN model

Chapter 3

Proposed Method

In this chapter, we introduce our proposed method for the availability improvement of IGMP proxy server. At the beginning, we state the requirements and assumptions of this research. Then, we explain our unique methods in details, such as:

- System architecture design
- Join, general query, leave, failover, and resuming flow design
- Fast redirection mechanism
- System VLAN design
- Fast standby synchronization mechanism

3.1 Requirements

The objective of the work is to design a high-availability IGMP Proxy server, which is suitable to be embedded in an access equipment, such as Optical Line Terminal (OLT) or DSLAM. However, the following descriptions about my proposed method will adopt Ethernet Passive Optical Network (EPON) OLT as an example.

In this work, some requirements as shown below should be specified before design:

- Hardware Platform Requirements

- High availability hardware platform: Motorola ATCA Chassis and Switch blade (F101)
- Hardware platform topology: Dual-star

- Functional and Performance Requirements
    - Support Simple & Fast Failover and Seamless Resuming
    - Support IGMP Proxy Redundancy Scheme: 1+1 (Active-Active)
    - Support IGMP Proxy Load Balance
    - Support dedicated Service VLAN (S-VLAN) (or Multicast VLAN (MC-VLAN)) for IPTV Service

3.2 System Architecture Design

Based on the above design requirements, a dual-star configuration is adopted as the topology of the platform architecture. Dual-star is a centralized-switch architecture as shown in Fig. 3-1. There are two switch blades as the central role of the whole ATCA system. These two switch blades are usually also as an interface with core network for the system. The switch blade is a Layer 2+ switch essentially, which is responsible to exchange data traffic among FRUs (Field Replaceable Unit) and uplinks. In other words, all outgoing and incoming data traffic has to through switch blades. Usually, the switch blades are not only for data plane exchanging, but also for control plane.

An EPON OLT system is an edge aggregation equipment which is always a tree topology has only one connection to the core network side and has many connections to the customer side. As describe in RFC 4605, an IGMP Proxy server is suitable to reside at the root of the tree, while it is easily to learn and proxy group membership information and simply forward multicast packets based upon that information. Based on the reason, we propose to put IGMP Proxy server function on switch blades of the ATCA-based EPON OLT system.

For high availability purpose, IGMP Proxy server function resides on both switch blades, and each switch blade has its individual path connecting to IPTV network through a dedicated multicast router as shown in Fig. 3-1.

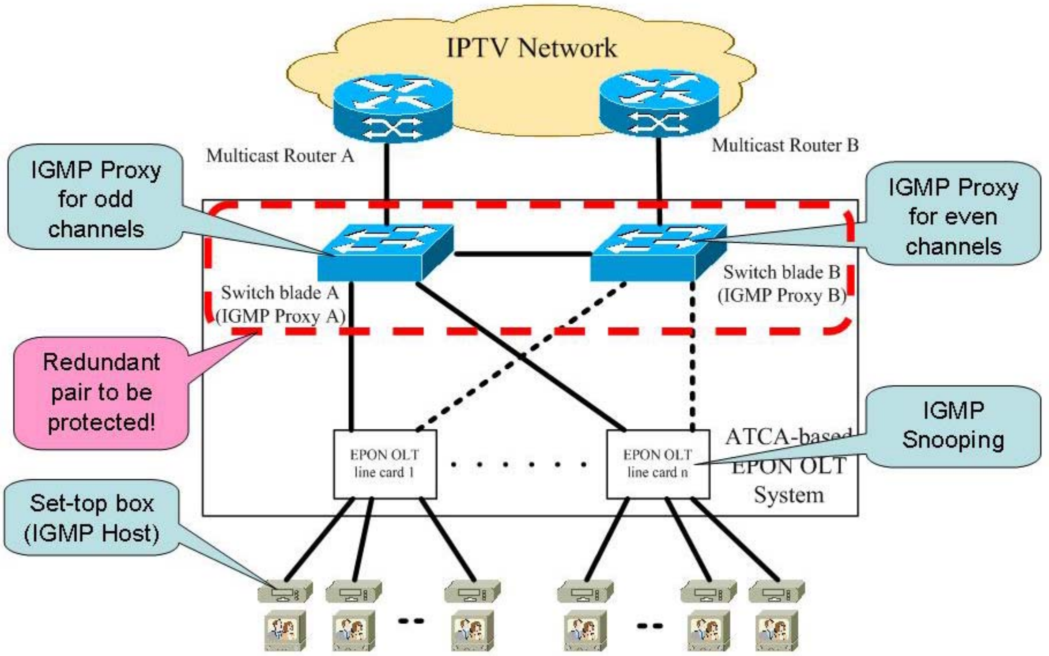

IPTV Network
Multicast Router A
Multicast Router B
IGMP Proxy for odd channels
IGMP Proxy for even channels
Switch blade A (IGMP Proxy A)
Switch blade B (IGMP Proxy B)
Redundant pair to be protected!
IGMP Snooping
ATCA-based EPON OLT System
EPON OLT line card 1
EPON OLT line card n
Set-top box (IGMP Host)

Fig. 3-1 ATCA-based EPON OLT system network architecture for IPTV service

Fig. 3-2 shows the architecture of ATCA-based EPON OLT system in details. An ATCA-base EPON OLT system can support up to 12 EPON OLT line cards, which connect with many customers. Each line card has two kinds of channels, Fabric and Base, to both of switch blades. Generally, Fabric channel is used for data plane, and Base channel is for control plane. In the work, the IPTV multicast traffic and IGMP message are through the Fabric channel, and the Health-Link-Check message through

Base. Due to high availability purpose, each line card should have separate paths to each switch blade. It guarantees at least one available path for each line card even if anyone of switch blades is failed at anytime.

Fig. 3-2 also shows the architecture of switch blade in details. The switch blade is a Layer 2+ switch essentially. It is mainly composed of Fabric interface switch system, Base interface switch system, RTM, and CPU. The Fabric interface and Base interface switch systems are isolated in principle. CPU is the only intersection between these two interfaces and possible to receive all traffic from both with right policy configuration on both Fabric and Base switch system. RTM, in this work, is used for the uplink interface to connect to multicast router in core network.

As describe in Chapter 1, IGMP Proxy server function is implemented on switch blades of an ATCA-based EPON OLT system. To say more precisely, the IGMP proxy server function is implemented on the CPU of switch blade. Any Multicast Join and Leave message from line cards are sent to CPU (IGMP Proxy server) of both switch blades through Fabric channel. The IGMP Proxy server configures the Fabric interface switch system with right VLAN, filtering settings and forwarding policies based on the above multicast membership information. Because of the consideration of system efficiency, the multicast traffic (e.g. video streaming) from multicast router will not be forwarded to and processed by CPU. The multicast traffic duplication and forwarding jobs are done by the Fabric interface switch system of switch blade.

The Base interface switch system is used to maintain health link check mechanism among two switch blades and all line cards. The healthy information of switch blade is broadcasted to all line cards though the base channel. The information is important and used to decide the uplink fabric path for a line card. There is also an

Update channel which is used for health link check and initial synchronization between the two switch blades.

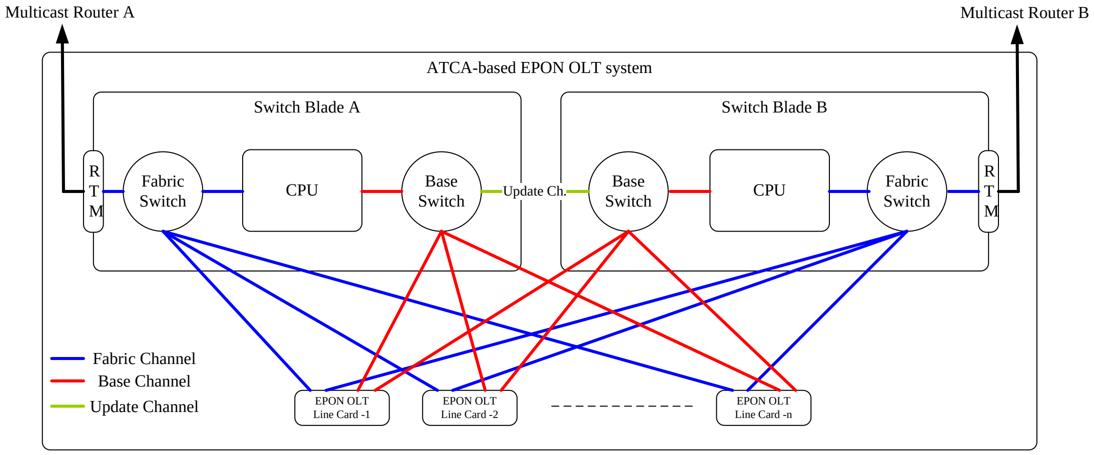


Fig. 3-2 ATCA-based EPON OLT system architecture

The system model from the view of Availability Management Framework (AMF) of Application Interface Specification (AIS) of Service Availability Forum (SAF) is illustrated in Fig. 3-3.

In my design, there is one service group (SG1), which is distributed between the two switch blades. SG1 is composed of two IGMP Proxy service units (S1 and S2), which provide IGMP Proxy service for odd and even channel in normal condition respectively, and which also reside on the two switch blades respectively.

In each service unit, there are two components (IGMP Proxy protocol service and Ethernet switch service). As the name of the components shown, IGMP Proxy protocol service components (C1 and C3) are responsible for handling the service of IGMP proxy, and the Ethernet switch service (C2 and C4) are for the functionality of Ethernet switch, which supports Multicast and VLAN especially. The workload of the both service units are assigned active HA state in a service instance, which is a logical entity for the aggregation of component service instances. Any failure occurring in any component within an active service unit will causes the entire service unit and all components within the service unit to be withdrawn from service.

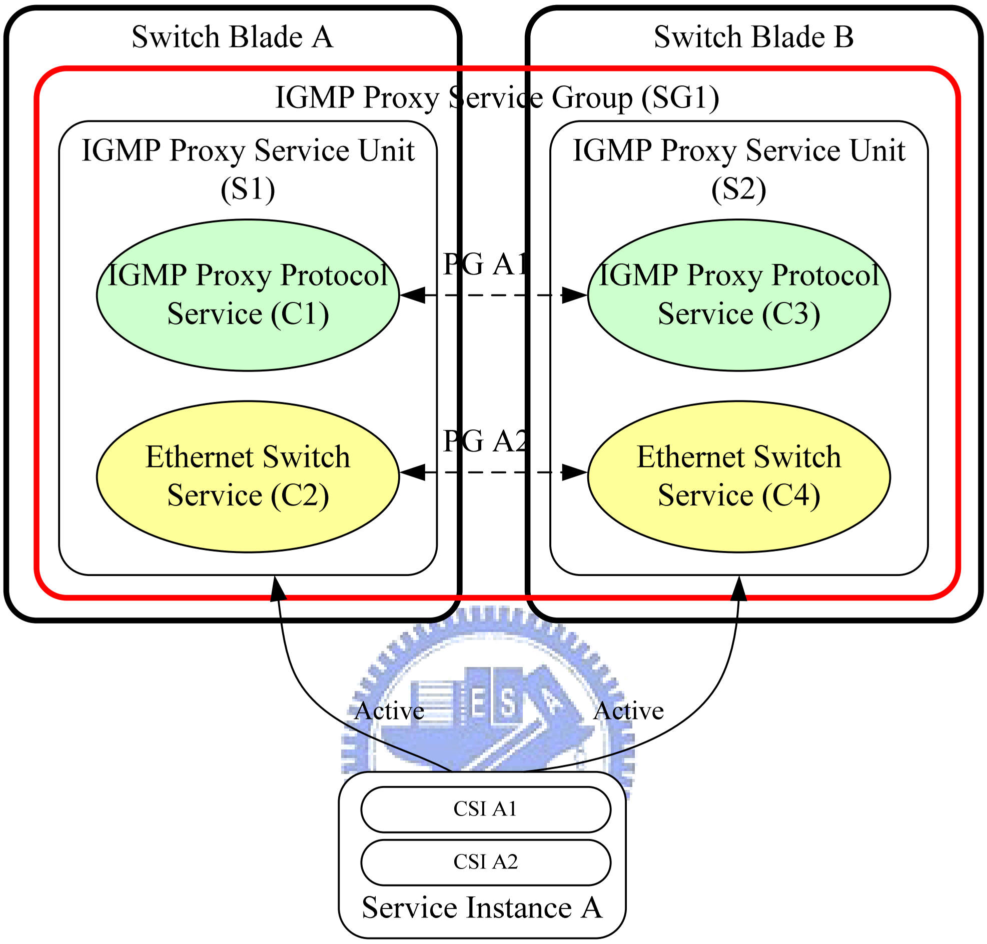

Switch Blade A
Switch Blade B
IGMP Proxy Service Group (SG1)
IGMP Proxy Service Unit (S1)
IGMP Proxy Service Unit (S2)
IGMP Proxy Protocol Service (C1)
IGMP Proxy Protocol Service (C3)
PG A1
Ethernet Switch Service (C2)
Ethernet Switch Service (C4)
PG A2
Active
Active
CSI A1
CSI A2
Service Instance A

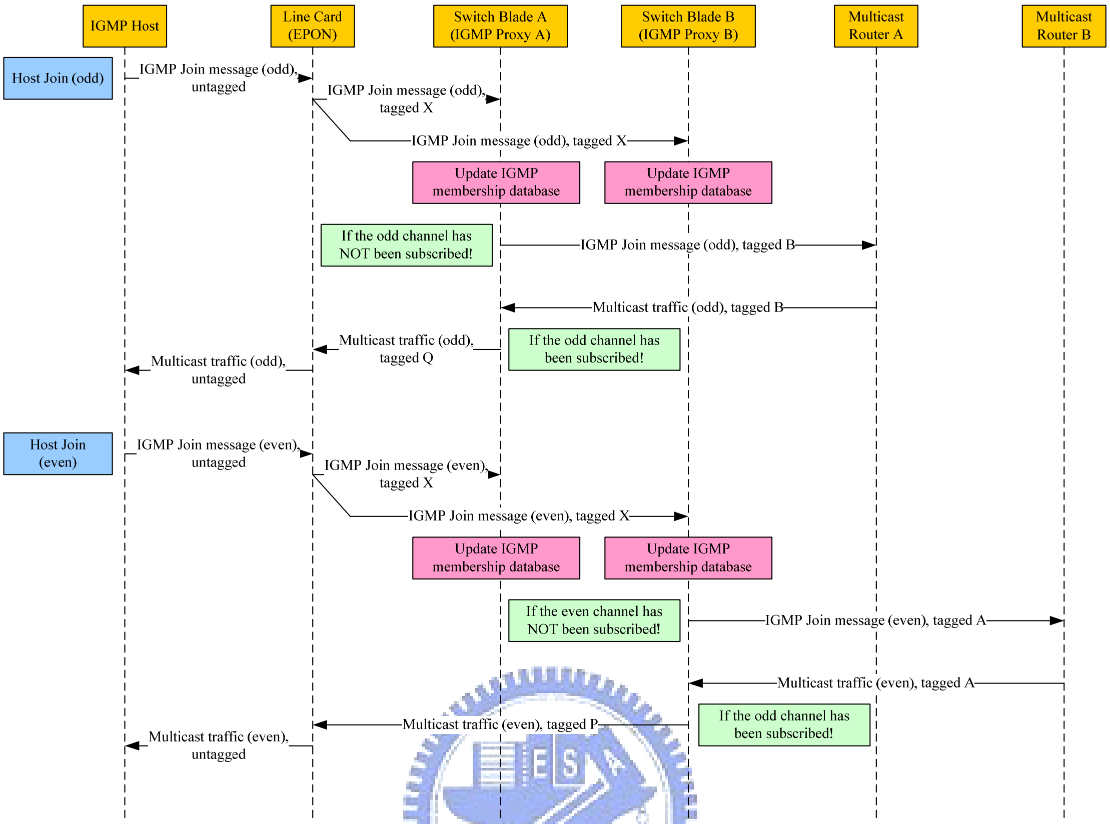
Fig. 3-3 Elements of system model of IGMP proxy service

3.3 Operational Scenario

In this chapter, we organize the following five operational scenarios (join, general query, leave, failover, and resuming) for a Carrier-Grade IGMP Proxy Server.

3.3.1 Join Scenario

The following sequence diagram shows the overall operations about that an IGMP host indicates that it wants to receive the specific multicast channel traffic (odd/even) from (“become a member of”) a multicast group.



IGMP Host
Line Card (EPON)
Switch Blade A (IGMP Proxy A)
Switch Blade B (IGMP Proxy B)
Multicast Router A
Multicast Router B
Host Join (odd)
IGMP Join message (odd), untagged
IGMP Join message (odd), tagged X
IGMP Join message (odd), tagged X
Update IGMP membership database
Update IGMP membership database
If the odd channel has NOT been subscribed!
IGMP Join message (odd), tagged B
Multicast traffic (odd), tagged B
Multicast traffic (odd), tagged Q
If the odd channel has been subscribed!
Multicast traffic (odd), untagged
Host Join (even)
IGMP Join message (even), untagged
IGMP Join message (even), tagged X
IGMP Join message (even), tagged X
Update IGMP membership database
Update IGMP membership database
If the even channel has NOT been subscribed!
IGMP Join message (even), tagged A
Multicast traffic (even), tagged A
Multicast traffic (even), tagged P
If the odd channel has been subscribed!
Multicast traffic (even), untagged

Fig. 3-4 Sequence diagram of join scenario

3.3.2 General Query Scenario

The following sequence diagram shows the overall operations about that an IGMP router (and an IGMP Proxy server) can ask the hosts which groups they are members of.

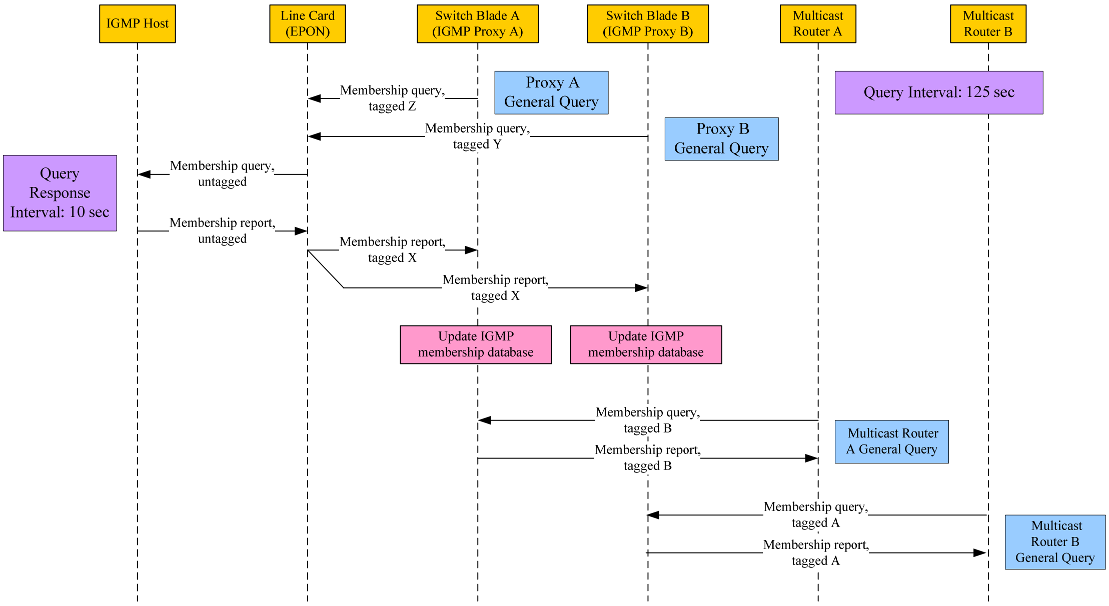

IGMP Host
Line Card (EPON)
Switch Blade A (IGMP Proxy A)
Switch Blade B (IGMP Proxy B)
Multicast Router A
Multicast Router B
Proxy A General Query
Query Interval: 125 sec
Membership query, tagged Z
Membership query, tagged Y
Proxy B General Query
Query Response Interval: 10 sec
Membership query, untagged
Membership report, untagged
Membership report, tagged X
Membership report, tagged X
Update IGMP membership database
Update IGMP membership database
Membership query, tagged B
Multicast Router A General Query
Membership report, tagged B
Membership query, tagged A
Multicast Router B General Query
Membership report, tagged A

Fig. 3-5 Sequence diagram of general query scenario

3.3.3 Leave Scenario

The following sequence diagram shows the overall operations about that an IGMP host indicates that it no longer wishes to receive the specific multicast channel traffic from a multicast group.

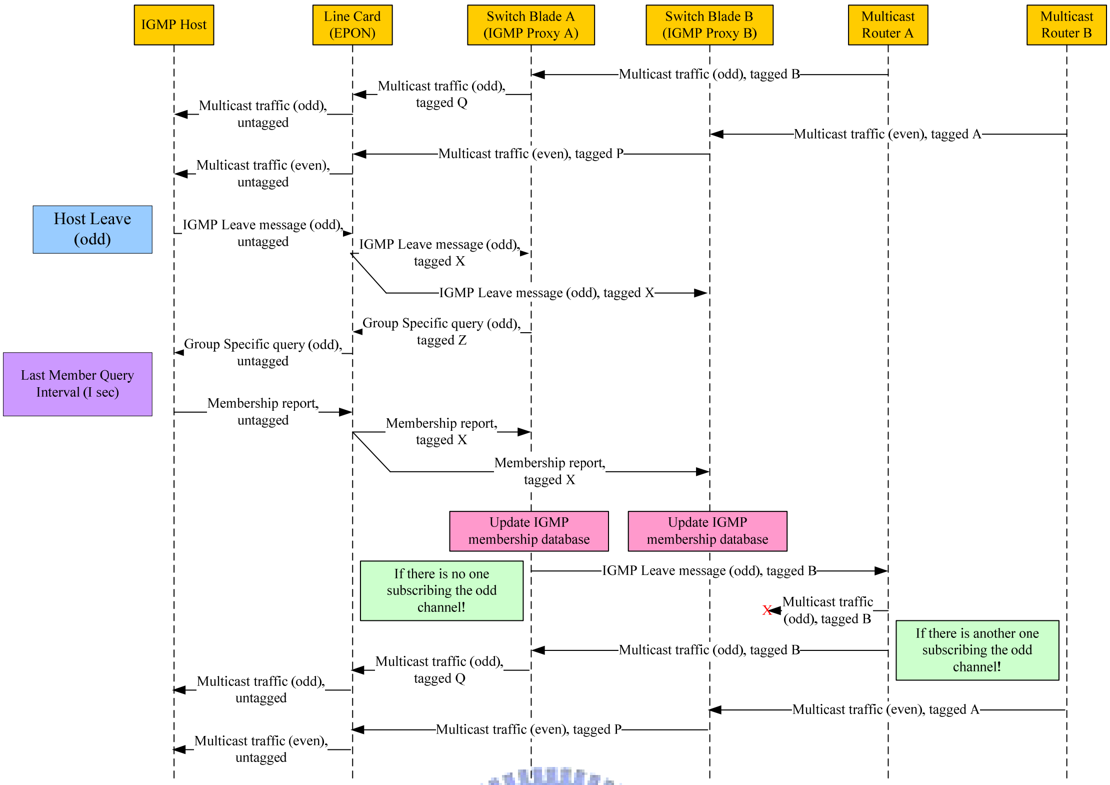
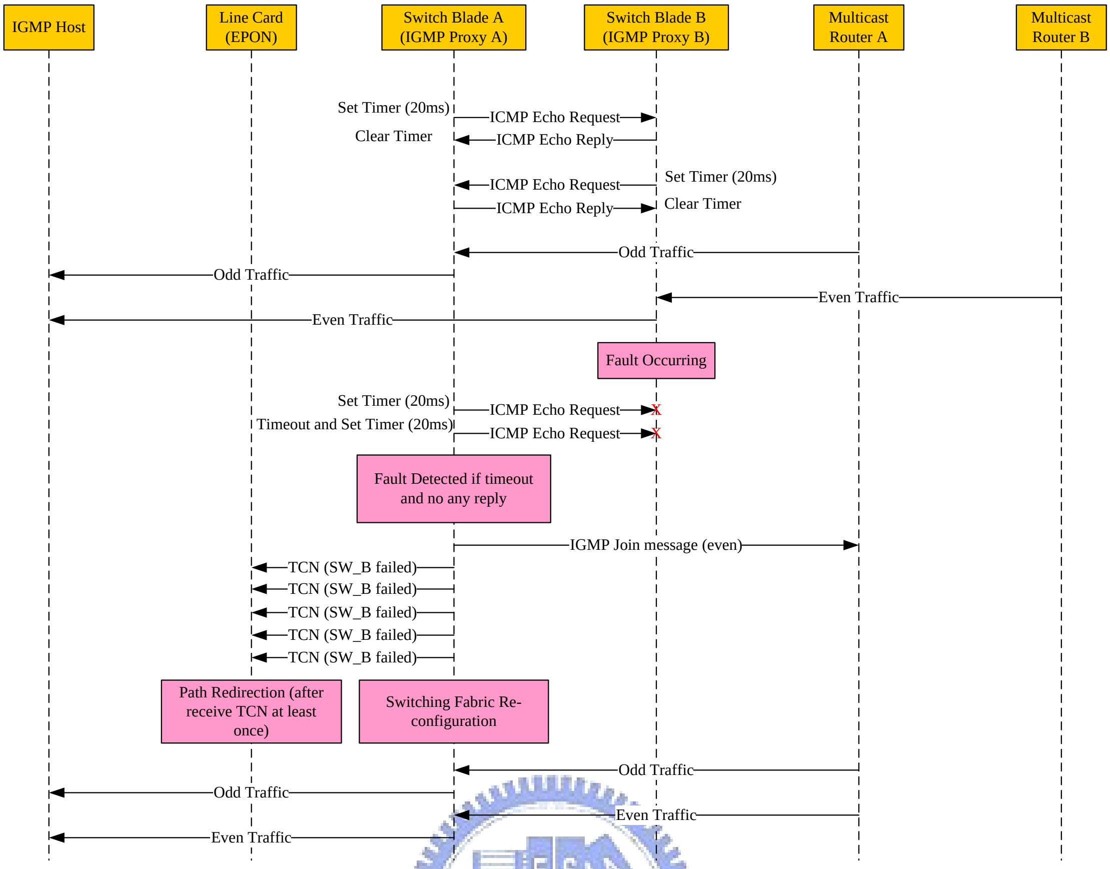

IGMP Host
Line Card (EPON)
Switch Blade A (IGMP Proxy A)
Switch Blade B (IGMP Proxy B)
Multicast Router A
Multicast Router B
Multicast traffic (odd), tagged B
Multicast traffic (odd), tagged Q
Multicast traffic (odd), untagged
Multicast traffic (even), tagged A
Multicast traffic (even), tagged P
Multicast traffic (even), untagged
Host Leave (odd)
IGMP Leave message (odd), untagged
IGMP Leave message (odd), tagged X
IGMP Leave message (odd), tagged X
Group Specific query (odd), tagged Z
Group Specific query (odd), untagged
Last Member Query Interval (I sec)
Membership report, untagged
Membership report, tagged X
Membership report, tagged X
Update IGMP membership database
Update IGMP membership database
If there is no one subscribing the odd channel!
IGMP Leave message (odd), tagged B
X
Multicast traffic (odd), tagged B
If there is another one subscribing the odd channel!
Multicast traffic (odd), tagged B
Multicast traffic (odd), tagged Q
Multicast traffic (odd), untagged
Multicast traffic (even), tagged A
Multicast traffic (even), tagged P
Multicast traffic (even), untagged

Fig. 3-6 Sequence diagram of leave scenario

3.3.4 Failover Scenario

The following sequence diagram shows the overall operations about while one of IGMP Proxy server failed, the others immediately take over the job of the failed one.



TCN: Topology Change Notification (Multicast)

Fig. 3-7 Sequence diagram of failover scenario

3.3.5 Resuming Scenario

The following sequence diagram shows the overall operations about while a new IGMP Proxy server resuming, the existing one seamlessly release the original job to the new one.

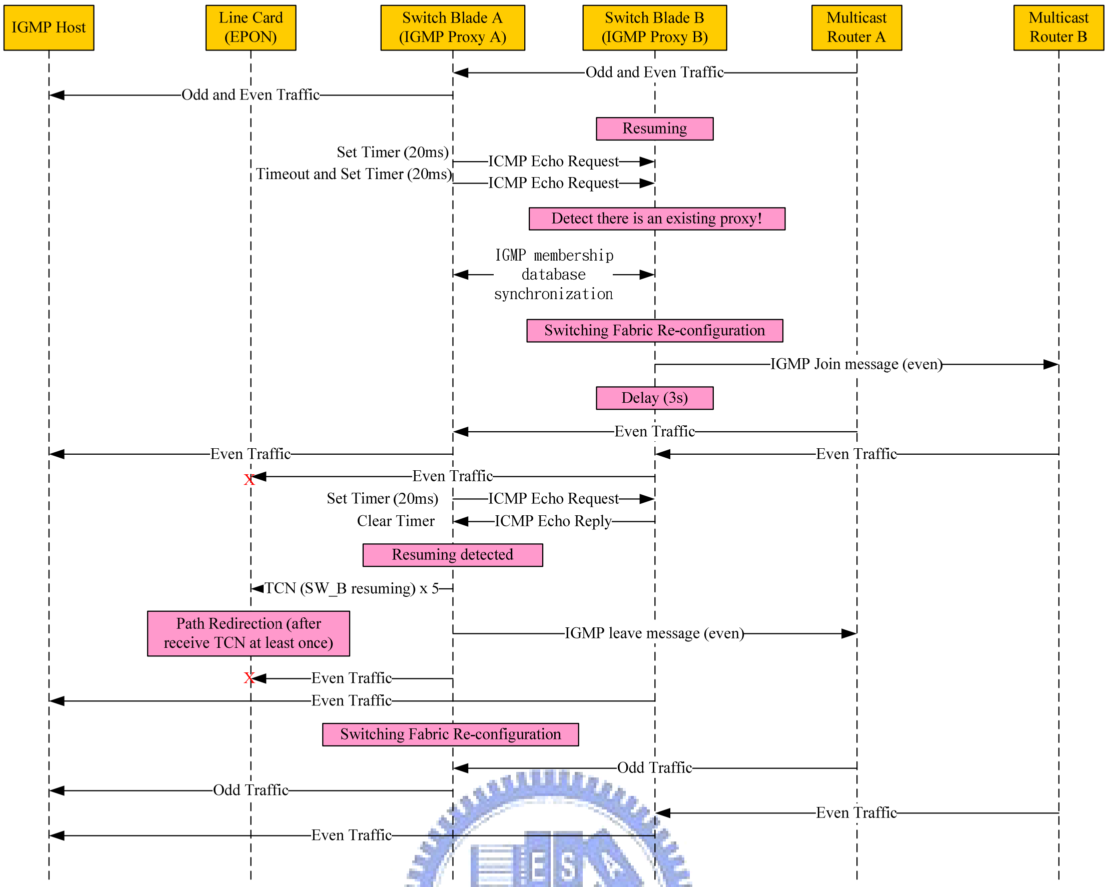


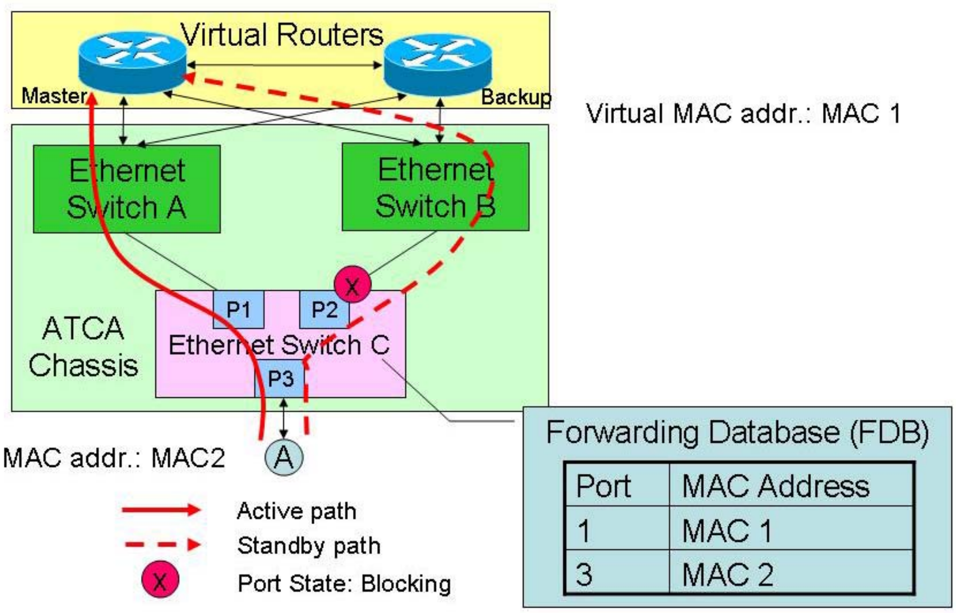
Fig. 3-8 Sequence diagram of resuming scenario

3.4 Fast Redirection

In this chapter, we will propose a method to solve the flushing of Forwarding Database (FDB) at the failover transition.

Fig. 3-9 shows the ideal case of failover transition. For normal operation, “A” sends a frame to Virtual Router (VR) through the active path (P3 => P1 => Switch A => VR) and the standby one is disable by blocking P2 of Ethernet Switch C. If there is a failure condition on Switch A, the P2 is enabled immediately and the standby path (P3 => P2 => Switch B => VR) is changed to the active one.

In real case, each Ethernet Switch transmits frames to a learned port recorded on the FDB. If a Switch transmits frames based on the FDB created before a failure has occurred, the frames cannot reach the destination, because the FDB is not updated after the failure (Fig. 3-10). Therefore, in all current technologies, when a failure occurs, the information of the FDB should be flushed. In order to flush the FDB, the MAC address must be deleted from FDB entries, and re-learned and restored as FDB entries. If there are many entries in FDB, a flushing takes a longer time. It will result in a longer failover time.



Port	MAC Address
1	MAC 1
3	MAC 2

Fig. 3-9 Ideal case of upstream redirection

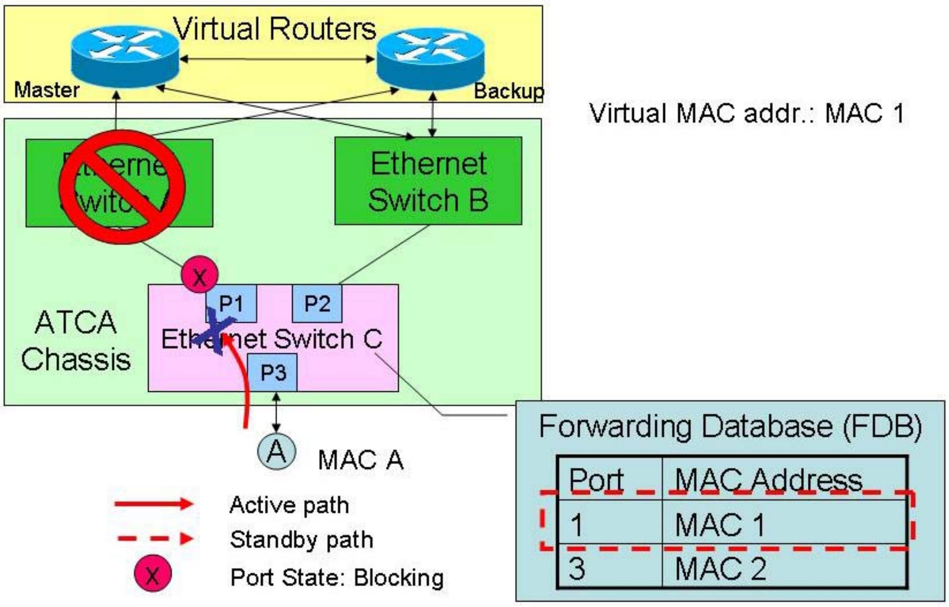


Fig. 3-10 Upstream redirection problem if without flush FDB

I propose a method using port trunking or aggregation method to eliminate the flushing of FDB problem at the failover transition. As shown in Fig. 3-11, if P1 and P2 ports of Ethernet Switch C are aggregated as a logic port T1, the entry in FDB will have no difference before and after a failure. In other words, we do not need to flush FDB at the failover transition.

Fig. 3-12 shows a derivative architecture using above port aggregation concept for Active-Active configuration.

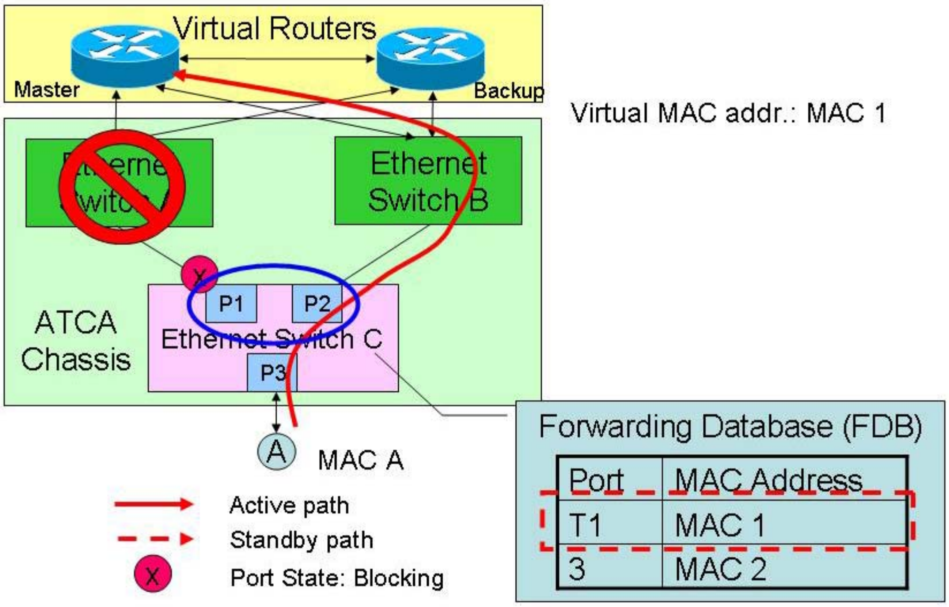


Fig. 3-11 Upstream redirection using port aggregation

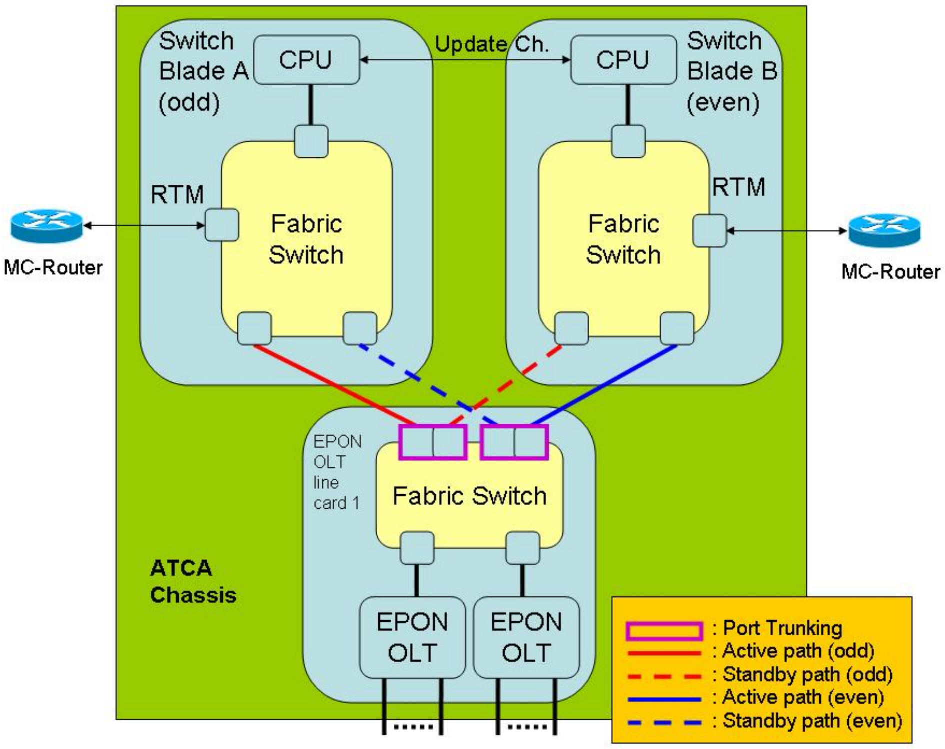


Fig. 3-12 Architecture using port aggregation method for active-active configuration

3.5 VLAN Design

In the section, we will explain how to plan the VLAN scheme to achieve load sharing for IPTV service and to introduce into C-VLAN and MC-VLAN in the ATCA-based EPON system for IPTV service.

In my research, we adopt VLAN mechanism for the following goals:

- To avoid packet loop: Need no STP
- To support IGMP Proxy Load Balance
  - Switch blade A (IGMP Proxy server A): IPTV odd channels
  - Switch blade B (IGMP Proxy server B): IPTV even channels.
  - The odd and even channels are identified according to the LSB of the multicast IP address.
- To support Multicast VLAN (Shared VLAN) for IPTV Service
- To limit the broadcasting scope of IGMP message and multicast traffic and to separate the odd- and even- channel IPTV traffic

For load sharing purpose, we proposed to divide the IPTV channels into two groups (odd and even channels) in Fig. 3-12. In normal case, the switch blade A (IGMP Proxy server A) is responsible for handling the odd channels, and the switch blade B (IGMP Proxy server B) for the even channels. The odd and even channels are identified according to the LSB of the multicast IP address. Assume the IPTV channels requested to deliver in normal distribution. The grouping proposal is sharing the load fairly between the two IGMP Proxy servers.

Due to IGMP message and multicast traffic will broadcast within the switch-based system, if without right configurations. To separate the odd- and even- channel IPTV traffic and to limit the broadcasting scope in a system, we adopt VLAN mechanism. The following table shows the VLAN usage in my design.

VLAN	Functions
VLAN X	Internal VLAN for IGMP Join/Leave from hosts to both proxies
VLAN Y	Internal VLAN for IGMP Query form even proxy server to hosts
VLAN Z	Internal VLAN for IGMP Query form odd proxy server to hosts
VLAN A	External VLAN for Multicast even-channel service (MC-VLAN)
VLAN B	External VLAN Multicast odd-channel service (MC-VLAN)
VLAN P	Internal VLAN for Multicast even-channel traffic
VLAN Q	Internal VLAN for Multicast odd-channel traffic

Table 3-1 VLAN list

The following sections give explanations in details for normal case (without failover) about the VLAN design individually.

3.5.1 Internal VLAN for IGMP Join/Leave from Hosts to Both Proxies

VLAN X is used to limit the broadcasting scope of IGMP join and leave message, which is from hosts to both IGMP proxies. An IGMP join or leave message from a host is duplicated by the switching fabric in the line card and broadcasted to both IGMP proxies. Both of IGMP proxies will receive the same message at the same time. This is an important point to realize the synchronization mechanism between both IGMP proxies.

Besides above, to prevent from a loop on the fabric switch of the switch blade, a right forwarding rule should be applied. It has to guarantee never to forward packet from downlink interface to downlink interface. All multicast traffic with VLAN X from downlink interface can only be forwarded to the CPU attached port.

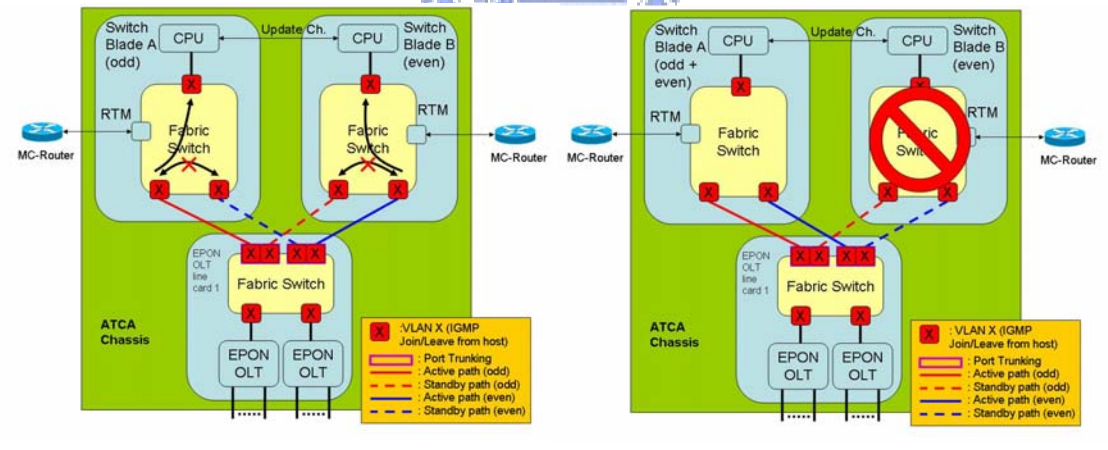


Fig. 3-13 Internal VLAN for IGMP join/leave from hosts to both proxies

3.5.2 Internal VLAN for IGMP Query form Proxies to Hosts

VLAN Y and Z are used to limit the broadcasting scope of IGMP query message, which is from even- and odd-channel IGMP proxies to all hosts. The even- and odd-channel IGMP queries are separated.

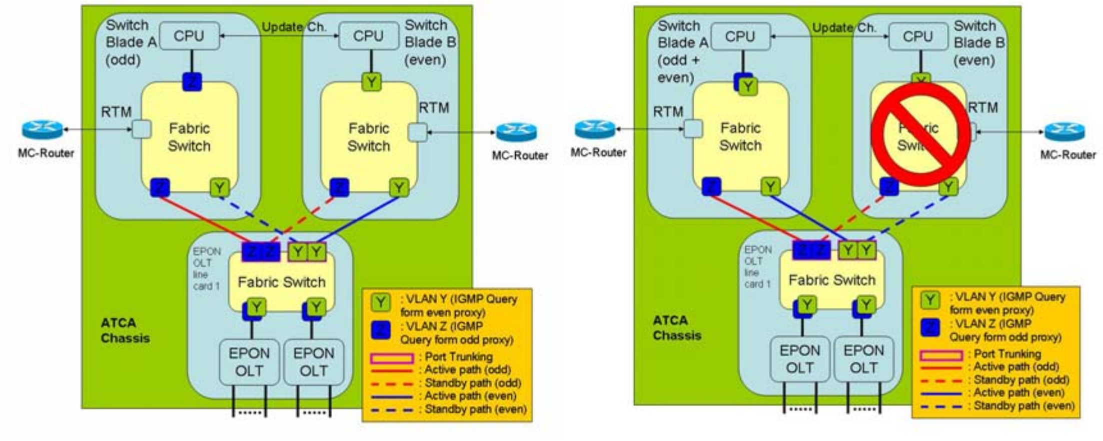


Fig. 3-14 Internal VLAN for IGMP query form proxies to hosts

3.5.3 External VLAN for Multicast Service (MC-VLAN)

VLAN A and B are used to limit the broadcasting scope of IGMP join/leave/query message, which is between IGMP proxies and external multicast routers. The VLAN A and B are just Multicast VLAN (MC-VLAN) represented even- and odd- channel IPTV service individually in my design. In other words, every join/leave/query message and multicast traffic from/to multicast router on the external interface is always tagged with the MC-VLAN.

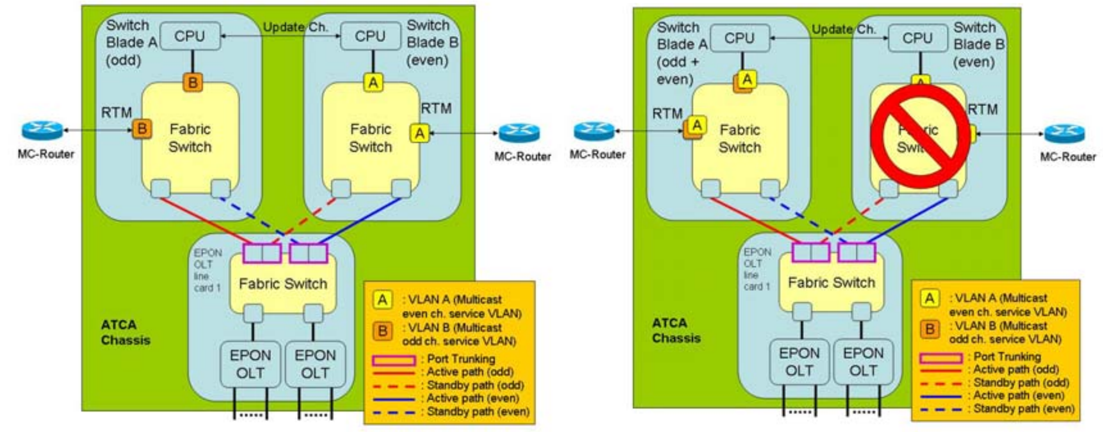


Fig. 3-15 External VLAN for multicast service

3.5.4 Internal VLAN for Multicast Traffic

VLAN P and Q are used to limit the broadcasting scope of multicast traffic (e.g. video streaming), which is from the RTM to all subscriber hosts within the system. The RTM is the port, which connects to external multicast router.

In order to prevent the multicast traffic from being forwarded to CPU and to follow the design concept of MC-VLAN, a right VLAN-change rule should be applied on the RTM port of the fabric switch. The rule guarantees to change the VID from MC-VLAN (VLAN A and B) to VLAN P and Q for all incoming multicast traffic (UDP).

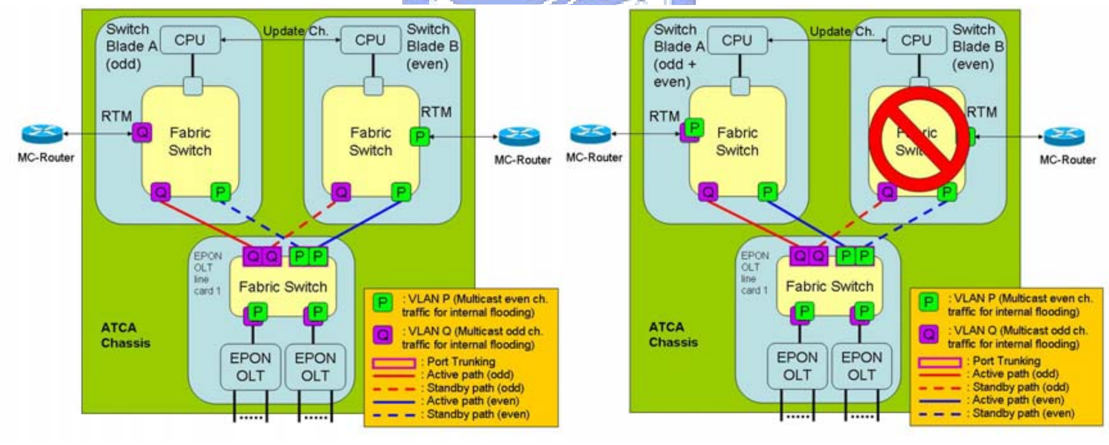


Fig. 3-16 Internal VLAN for multicast traffic

3.6 Standby Synchronization

How to synchronize between a redundant pair of IGMP Proxy Servers? The synchronization mechanism is an important issue, which greatly affects failover performance, system complexity, and CPU loading. we propose a simple way to use the natural characteristic of multicast to achieve synchronization between a redundant pair.

The IGMP proxy device has to maintain a database consisting of the merger of all subscriptions on any downstream interface. The IGMP proxy device sends IGMP membership reports on the upstream interface when queried and sends unsolicited reports or leaves when the membership database changes. The membership database is also used to decide the multicast traffic forwarding policies on the downstream interface. The content of membership database greatly affects the behavior of IGMP proxy server obviously. In other words, the same content of membership database of two IGMP proxies will have the same behavior.

For redundancy to work, the standby unit needs to be kept synchronized with the active unit at all times. This is required so that the standby can quickly take over the jobs of active unit in case the active one fails. In the work, Switch blade A (or IGMP proxy server A) serves the odd-channel IPTV service primary, and is a standby unit for even-channel service. Switch blade B (or IGMP proxy server B) is primary for even channel, and is standby for odd. To make the membership database of the two IGMP proxy servers synchronous for each others and for both odd and even IPTV services will have the best failover performance.

As describe above, the content of membership database is constructed according to the join and leave behavior on downstream interface. Due to a multicast packet (e.g. IGMP join or leave message) is naturally broadcast in an Ethernet-based switch system by default. My proposed method is to use the natural characteristic to achieve the membership database synchronization between a redundant pair of IGMP Proxy Servers. In Fig. 3-14, the IGMP join and leave message from the subscriber is broadcasted to both IGMP proxy servers by the switching fabric of EPON OLT line card within VLAN X. The advantage of the method guarantees both IGMP proxy servers receiving the same IGMP join and leave message at the same time. This is naturally done the synchronization between the redundancy pair.

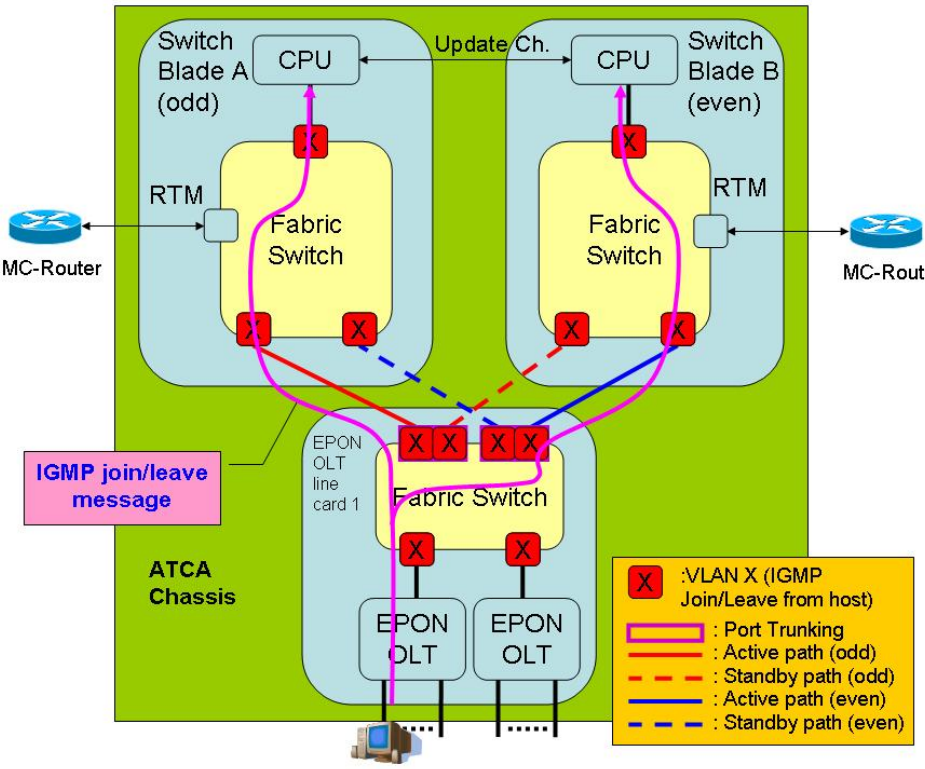


Fig. 3-17 IGMP join/leave message flow

My proposed method for standby synchronization should also cooperate with the following mechanisms to perfect the design.

- The switching fabric of both Switch blades should be configured allowing the IGMP join and leave messages only from downstream interface to upstream, not from downstream interface to downstream. This configuration can avoid a multicast storm of the IGMP join and leave message within the system.

- My proposed method for standby synchronization is like Lockstep mechanism. It relied on the redundant pair always receiving the same message to result in the same content in membership database. The synchronization mechanism is feasible only when the both IGMP Proxy Servers of the redundant pair have the same initial state. In fact, both IGMP proxy servers are usually not startup at the same time. Therefore, an extra synchronization procedure should be done in the initial state, shown as Fig. 3-15. My proposed method is using the Update channel between both IGMP Proxy Servers to do this.

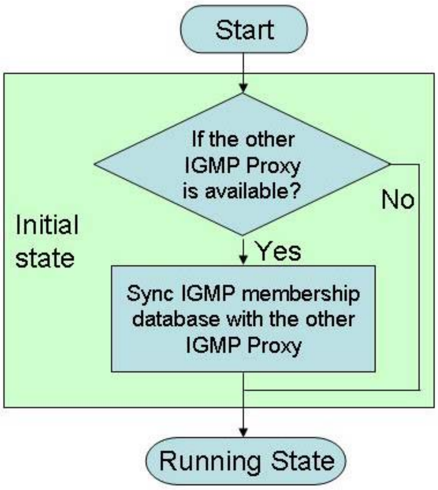


Fig. 3-18 Flow chart of standby synchronization in initial state

Chapter 4

Test and Results

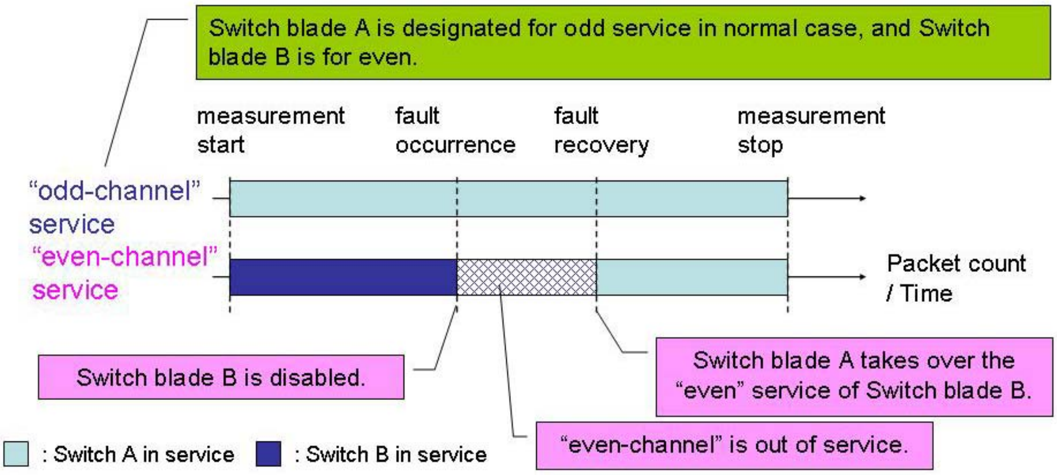
Failover time is the key performance index to judge a failover mechanism. In this chapter, we introduce the principle of failover time measurement, and our measurement environments and results.

4.1 Principle of Failover Time Measurement

Failover time is defined the amount of time it takes for Failover to successfully complete by Internet Engineering Task Force (IETF) [25]. It is the key performance index to judge a failover mechanism. Failover Time can be calculated using the Time-Based Loss Method (TBLM), Packet-Loss Based Method (PLBM), or Timestamp-Based Method (TBM). In the thesis, we adopt PLBM to measure the failover time.

Fig. 4-1 shows my measurement concept. First, we generate two identical testing flows with the same content, speed, and frame length, and with simultaneous start and stop operations. These two flows are named “odd-channel” and “even-channel” respectively and served by Switch blade A and B respectively in normal case. In the case without any failures, Switch blade A and B can serve the same amount of traffic within a period. If a failure condition occurs on Switch blade B, the "even-channel" service will be out of service for a short period, then the service is resumed after Switch blade A taking it over. However, the "odd-channel" is still running without any interruptions at that moment. It is clear that the difference of the amount of packet between "odd-channel" and "even-channel" can be easily transformed into failover time. The formula is shown as below.

$$\text{failover time} = \frac{\left(\text{RCVpacket}_{\text{odd}} - \text{RCVpacket}_{\text{even}}\right) \text{ x packet\_length (byte) x 8 bit}}{\text{Data Rate (Mbps)}} \qquad (4\text{-}1)$$



Switch blade A is designated for odd service in normal case, and Switch blade B is for even.
measurement start
fault occurrence
fault recovery
measurement stop
"odd-channel" service
"even-channel" service
Packet count / Time
Switch blade B is disabled.
Switch blade A takes over the "even" service of Switch blade B.
"even-channel" is out of service.
: Switch A in service
: Switch B in service

Fig. 4-1 Principle of failover time measurement

4.2 Testing Environment of Failover Time Measurement

Fig. 4-2 depicts the testing environment of failover time measurement. The testing environment setup is following EPON network general configuration. we use Sprient SmartBits to act the both roles of Multicast routers and IGMP clients to generate and to receive the testing flows.

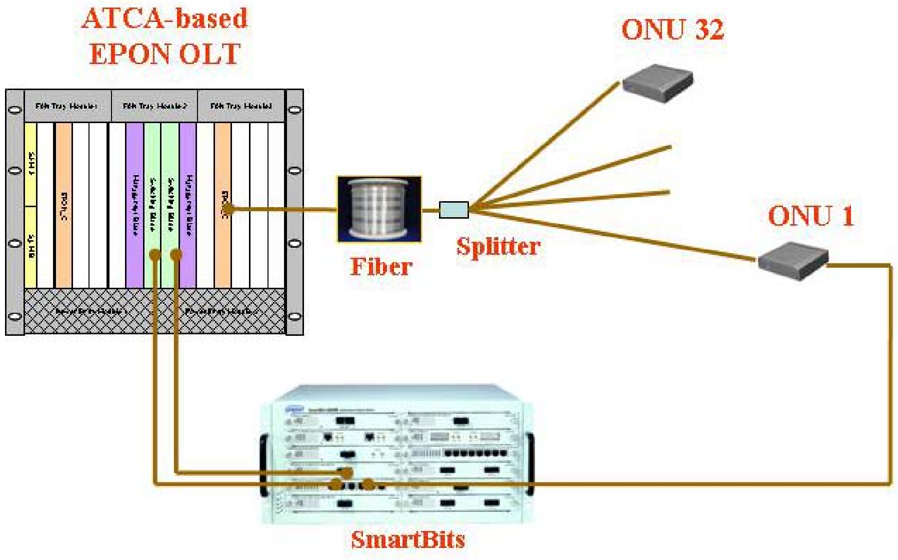


Fig. 4-2 The testing environment setup of failover time measurement

As Fig. 4-2 shown, SmartBits generates testing flow 1 and 2 to Switch blade A, and 3 and 4 to B. Both of the flow 1 and 3 are identical and represent the odd-channel multicast traffic. Both of the flow 2 and 4 are for the even-channel. Switch blade A only serves the odd-channel (flow 1) and Switch blade B serves the even-channel (flow 4) normally. If a failure condition occurs on Switch blade B, the even-channel service is interrupted for a short period. After that, Switch blade A takes over to serve the even-channel (flow 2) immediately. The total amount of served even-channel is the sum of received flow 2 and 4. The formula to calculate failover time is simplified as below.

$$\text{failover time} = \frac{\left(\text{flow 1} - \left(\text{flow 2} + \text{flow 4}\right)\right) \text{ x 128 byte x 8 bit}}{\text{50 Mbps}} \tag{4-2}$$

Where 128 byte is the packet length; 50Mbps is the data rate

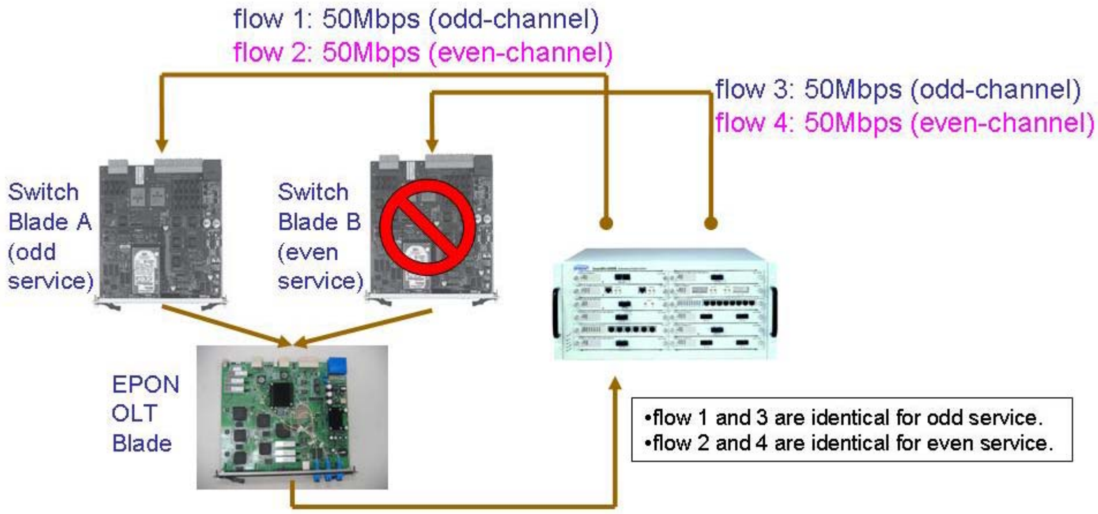

flow 1: 50Mbps (odd-channel)
flow 2: 50Mbps (even-channel)
flow 3: 50Mbps (odd-channel)
flow 4: 50Mbps (even-channel)
Switch Blade A (odd service)
Switch Blade B (even service)
EPON OLT Blade
•flow 1 and 3 are identical for odd service.
•flow 2 and 4 are identical for even service.

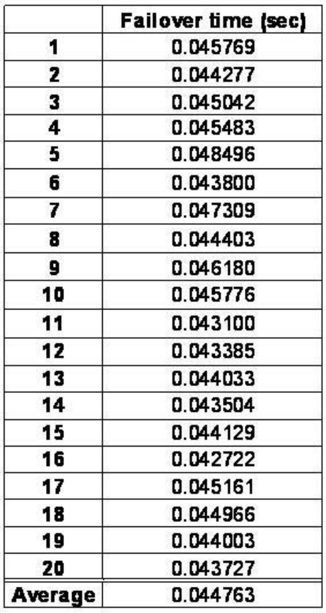
Fig. 4-3 The testing environment setup of failover time measurement

4.3 Results of Failover Time

In this paragraph, it shows both of the simulation and real results of failover time following the testing environment setup described in Chapter 4.2.

The simulation result is shown in Fig. 4-4. The average failover time is about 45ms. The performance is better than 50ms, which is the failover requirement specified by SONET/SDH.

The real result is shown in Fig. 4-5. The average failover time is around 400ms averagely. The result is much different with the simulation one. To further analyze the result, we found the failover time is mostly wasted by the VLAN reconfiguration via the opening SNMP interface on Motorola switch blade. It is an implementation issue does not be solved in this work.



	Failover time (sec)
1	0.045769
2	0.044277
3	0.045042
4	0.045483
5	0.048496
6	0.043800
7	0.047309
8	0.044403
9	0.046180
10	0.045776
11	0.043100
12	0.043385
13	0.044033
14	0.043504
15	0.044129
16	0.042722
17	0.045161
18	0.044966
19	0.044003
20	0.043727
Average	0.044763

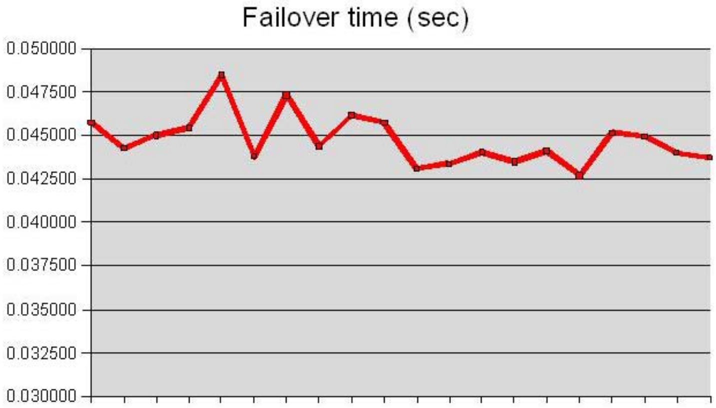


Fig. 4-4 The simulation results of failover time

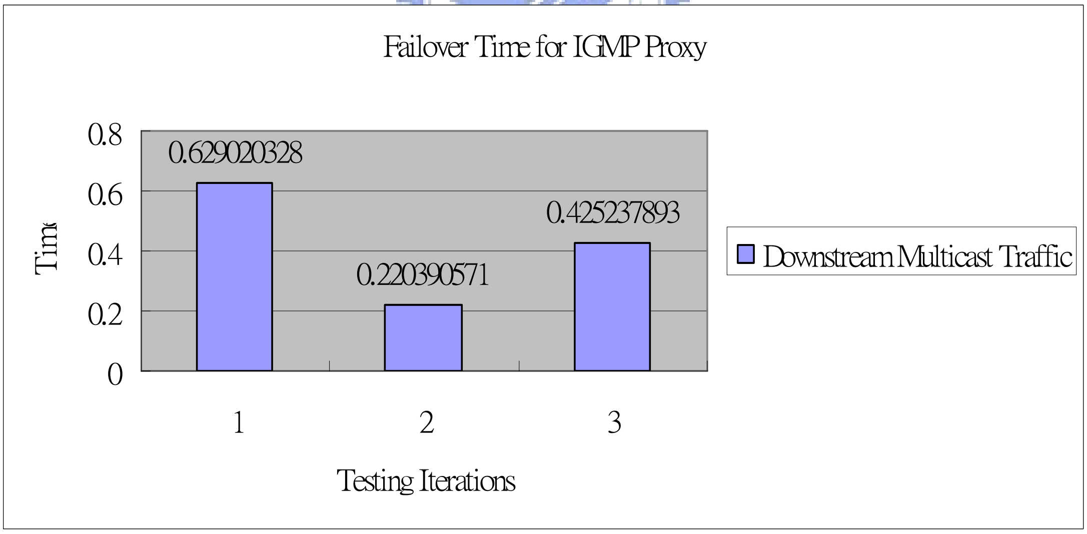


Fig. 4-5 The real results for failover time

Chapter 5

Conclusion

In this chapter, we make the conclusions and the future works for this research.

5.1 Conclusion

The objective of this research is to improve the availability of IGMP proxy function on an access equipment. In order to achieve the goal, we propose using redundant design with hot standby, load sharing, fast failover, and seamless resuming capabilities to reduce MTTR and to increase availability. In the research, we take the characteristics of VLAN, ATCA-based topology and multicast traffic to simplify the IGMP Proxy failover mechanism and to improve the performance. The following items are the distinguishing points of my proposed method:

- Simple VLAN configuration to enable active-active redundancy
- Simple standby synchronization mechanism for multicast packets
- Simple and fast redirection method to solve flushing FDB problem
- Simple seamless resuming procedure without interruption for IPTV service

In the thesis, my proposed architecture also supports the other advanced features, such as Multicast VLAN (Shared VLAN), for real IPTV Service deployment

The average simulation result of failover time can minimize to 45ms, which is better than the specifications of SONET/SDH.

5.2 Future Works

In the thesis, my proposed standby synchronization mechanism is suitable for all multicast packets, not only for IGMP. There are lots of Internet protocol, such as RIP2 and VRRP, also using multicast address. To apply the mechanism to the other IP protocols, which use multicast MAC address may be a worthy research.

As the description of Chapter 4.3, the VLAN configuration interface on Motorola Switch blade is slower. The platform issue results in the overall failover performance decreasing. To choose a better Switch blade may solve the problem.

In the thesis, we realized the IGMP Proxy following the IGMP protocol v2. It can be also upgraded to v3.

Bibliography

[1] B. Alfonsi, “I Want My IPTV: Internet Protocol Television Predicted a Winner,” IEEE Distributed Systems Online, IEEE Computer Society 1541-4922, vol. 6, no. 2, Feb.2005.

[2] HA Forum, “Providing Open Architecture High Availability Solutions”, Revision 1.0, February, 2001

[3] Peter S. Lau, “A best effort multicast algorithm,” submitted to ICCT’2006.

[4] W. Jia, W. Zhao, D. Xuan, and G. Xu,“ An Efficient Fault-Tolerant Multicast routing Protocol With Core-Based Tree Techniques,” IEEE Trans. On Parallel And Distributed Systems, Vol. 10, No. 10, October 1999.

[5] A. Fei, J. Cui, M. Gerla, and D. Cavendish,“ A Dual-Tree Scheme for Fault-Tolerant Multicast,” IEEE International Conference On Communications, 2001.

[6] Jun-Hong Cui, Michalis Faloutsos, and Mario Gerla, “An Architecture for Scalable, Efficient and Fast Fault-Tolerant Multicast Provisioning,” IEEE Network special issue on Protection, Restoration, and Disaster Recovery, 18(2), pp. 26-34, March-April, 2004.

[7] Muriel Medard, Steven Finn, Richard Barry, and Robert Gallager, “Redundant Trees for Preplanned Recovery in Arbitrary Vertex-Redundant or Edge-Redundant Graphs,” IEEE/ACM Trans. on Networking Vol. 7, No. 5, October 1999.

[8] A. Ballardie, “Core Base Tree (CBT) multicast routing architecture,” RFC2201, Sept. 1997.

[9] L. Schwiebert and R. Chintalapati, “Improved Fault Recovery for Core Based Tree,” Comp. Commun., vol. 23, no. 9, Apr. 2000.

[10] T. Pusateri, “Distance vector multicast routing protocol,” draft-ietf-idmr-dvmrp-v3-11, Oct. 22, 2003.

[11] Juan Wu, YaLing Nie, Hideya Yoshiuchi, Hiroki Ikeda, “Building Multicast Controller for Carrier-grade IPTV Service over Ethernet Passive Optical Network,” ICSNC, 2007

[12] “AdvancedTCA Base,” PICMG 3.0 R2.0, 18-Mar-05

[13] IEEE Computer Society, LAN MAN Standards Committee, “Media Access Control (MAC) Bridges,” IEEE 802.1d, 1998, <http://www.ieee802.org>.

[14] IEEE Computer Society, LAN MAN Standards Committee, “Media Access Control (MAC) Bridges,” IEEE 802.1d-2004, June 2004, <http://www.ieee802.org>.

[15] IEEE Computer Society, LAN MAN Standards Committee, “Resilient Packet Ring (RPR) access method and physical layer specifications,” IEEE 802.17, Sept. 2004, <http://www.ieee802.org>.

[16] Extreme Networks, “Extreme Standby Router Protocol and Virtual Routing Redundancy Protocol,” 2002, < http://www.imerja.com/files/file/White%20Papers/Extreme/VRRP%20vs%20ESRP.pdf >.

[17] Extreme Networks, “Ethernet Automatic Protection Switching (EAPS),” 2006, < http://www.extremenetworks.com/libraries/whitepapers/WEAPS_1293.pdf>.

[18] Elie Sfeir, Saiidrine Pasquahi, Thomas Schwabe, Andreas Iselt, “Performance Evaluation of Ethernet Resilience Mechanisms” , IEEE, 2005

[19] I. Van de Voorde, L. Tancevski, G. Chiruvolu, Y. T’ Joens, J. De Jaegher, “Carrier-grade Ethernet: extending Ethernet into next generation metro networks”, Alcatel Telecommunications Review - 3rd Quarter 2002, < http://lt.fe.uni-lj.si/gradiva/KOS/clanki_pdf/18-carrier%20grade%20ethernet%20-%20extending%20ethernet%20into%20next%20generation%20metro%20networks.pdf >

[20] Tim Hubbard, “Optimizing Metro Ethernet”, Nortel Networks, 2004, < http://metroethernetforum.org/PPT_Documents/IIR-MEF-Barcelona-11-04-Optimizing-MENs.ppt >

[21] “HIPAA Compliance: An Extreme Approach”,Extreme Networks, < http://www.forsitegroup.com/pdf/Extreme_HIPPA.pdf >

[22] S. Deering, “Host Extensions for IP Multicasting”, RFC 1112, August 1989

[23] W. Fenner, ” Internet Group Management Protocol, Version 2”, RFC 2236, November 1997

[24] B. Cain, S. Deering, I. Kouvelas, B. Fenner, A. Thyagarajan, “Internet Group Management Protocol, Version 3”, RFC 3376, October 2002

[25] S. Poretsky, R. Papneja, J. Karthik, S. Vapiwala, ” Benchmarking Terminology for Protection Performance”, IETF (draft), November 2007