PBT Networking

Host-to-host connections through SURFnet6

UvA Students
G.A. van Malenstein, B ICT
C. Steenbeek, B ICT

Supervisors
drs. R. van der Pol, SARA
A. Toonk, MSc, SARA
dr. ir. C. Th. A. M. de Laat, UvA

Research Project 2
System and Network Engineering

University of Amsterdam (UvA)

SARA

July 2, 2007

Abstract

SURFnet6 is a hybrid network through The Netherlands, linking universities and other research institutions together by high-speed optical lightpaths. Customers of SURFnet are able to create a lightpath, which delivers a reliable connection through SURFnet6 with Quality of Service (QoS). SARA came up with the idea to extend this QoS into the customers' networks, so a complete host-to-host connection could be made. This project is started, because the QoS is ended when entered the customer's network, at this moment. Provider Backbone Transport (PBT) is a new ethernet extension based on Provider Backbone Bridges, which enables layer 2 circuit-switched ethernet connections with QoS and reliability.

By describing the theoretical and possible practical operation of PBT and its related protocols, and researching multiple solutions in various types of networks, a solution for implementing PBT with SURFnet6 was found. When placing two or more PBT bridges in the customers' ethernet-based networks, it becomes possible to create an end-to-end connection with QoS and reliability through SURFnet6.

Acknowledgments

We would like to extend our gratitude to the following people and organizations for their assistance in our research:

SARA, Amsterdam – For facilitating their accommodation and assistance by all 'colleagues'.

drs. R. van der Pol, SARA – For supervising this research project and delivering input and insight on behalf of this project.

A. Toonk, MSc, SARA – For supervising this research project and delivering input and insight on behalf of this project.

Ir. G.W.J. Jacobs, Nortel Networks – For providing information about PBT and Nortel's investments in PBT.

dr. P. Grosso, UvA – For her input in this project and for discussing the interests of the University of Amsterdam.

Contents

Chapter 1

Introduction

On the 1st of January 2004, the Gigaport Next Generation project was launched in The Netherlands. This 5-year consortium consisting of Nortel, Avici, Telindus resulted in 2006 in the new SURFnet6 infrastructure, which is described in section 5.2. This optical network connects SURFnet's customers – colleges, research institutions and universities – to each other and to the internet.

SURFnet6 proves to be a fast and reliable network to date. With speeds up to ten gigabits per second, around 750,000 users are served on a daily basis with internet connectivity. Nowadays, scientists, who are connected to the net via SURFnet6, have the needs to set up host-to-host connections; making it possible to connect a research institute directly to a telescope at the other side of the world, for example [1].

In this report, the problem of creating host-to-host connections with a certain Quality of Service through SURFnet6 is being researched – see chapter 3 for the problem description. In chapter 2, the background of this problem is described. Chapter 4 describes how the research is carried out.

The topology of SURFnet6 and the techniques used to create and operate the network are described in chapter 5. In this report, the new ethernet extensions Provider Backbone Bridges and Provider Backbone Transport are described in chapters 6 and 7.

Research conducted on PBT in combination with SURFnet6 is covered in chapter 8. The conclusions and recommendations will be described in chapter 9.

Host-to-host ethernet connections can be made through SURFnet6 at the moment. However, SURFnet's customers all have ethernet-based networks, so it is impossible to guarantee a certain Quality of Service for these connections. The arisen problem will be described in the next chapter.

Chapter 2

Background

Ethernet technology exists as a communication protocol since 1973 and has proven its value in Local Area Networks (LANs). Today, the need for building ethernet-based networks throughout a large metropolitan area – Metropolitan Area Networks (MANs) – exists. Ethernet protocols are packet-switched instead of circuit-switched. Also, the Quality of Service (QoS) on ethernets cannot be guaranteed on certain circuits, because no OAM (Operations, administration, and maintenance) functionalities are built into ethernet protocols. With these features being absent, setting up an end-to-end protected path with QoS through an ethernet network is impossible. At this moment, new extensions to the ethernet protocols (802.1ag, 802.1ah, 802.1Qay) are being developed to transform ethernet to a technology ready for use in MANs [2].

SURFnet6 is a hybrid network owned by SURFnet, divided into a traditionally routed IP section, and a new innovative optical section. Via this optical section, lightpaths – described in section 5.1 – with speeds of one to ten gigabits per second are offered to customers. Customers of SURFnet have their own ethernet-based network, which is connected to other customers and the internet by the SURFnet6 infrastructure. When customers of SURFnet create host-to-host connections, the sections where ethernet is being used, offer no Quality of Service or protected paths. Monitoring of SURFnet6 and technical maintenance is done by a NOC consisting of SARA and Telindus.

The SURFnet website describes the organization as:

> "SURFnet is a subsidiary of the SURF organization, in which Dutch universities, universities for applied sciences and research centers collaborate nationally and internationally on innovative ICT facilities. Other subsidiaries of the SURF organization are SURFfoundation and SURFdiensten (SURFservices).[1]"

SARA presents itself as follows at its website:

> "SARA Computing and Networking Services is an advanced ICT service center that supplies since more than 30 years a complete package of high performance computing & visualization, high performance networking and infrastructure services. Among SARAs customers are the business community and scientific, educational, and government institutions.[2]"

The Telindus website provides the following company information:

> "Telindus is a group of companies offering ICT Solutions and Services on an international level. We serve business, service provider and public market needs. With 2.700 employees, Telindus is present in 14 countries in Western Europe, Sweden, Hungary, China and Thailand.[3]"

2.1 Carrier ethernet

In large telecom networks, OAM – explained in section 6.2.1 – is used widely. In next-generation networks as carrier ethernet, implementing OAM will be a challenge; ethernet exists at the moment without OAM.

[1] Source: SURFnet.nl
[2] Source: SARA.nl
[3] Source: Telindus.com

The term for ethernet services and the use of ethernet as a transport infrastructure on Metropolitan Area Networks and Wide Area Networks – instead of only LANs – is *carrier ethernet* [3]. Carrier ethernet is only ready for worldwide adoption, when OAM functionality is added to ethernet. The ethernet standards 802.1ag and Y.1731 are being developed to enable OAM on ethernet.

2.1.1 Requirements

> "Carrier Ethernet is a ubiquitous, standardized, carrier-class SERVICE defined by five attributes that distinguish Carrier Ethernet from familiar LAN based Ethernet [4]."

When creating a carrier ethernet based entirely on the ethernet protocols, availability of the network and a certain quality of the connections has to be assured. The extensions for monitoring are currently in development. The following description contains all requirements in order to set up an ethernet with these new extensions [4]:

Standardized Services No changes to the customer's LAN equipment are needed;

Quality of Service Many graduations in the bandwidth and service options are available. Based on Service Level Agreements (SLAs) that provide end-to-end performance, based on committed information rate, frame loss, delay and jitter;

Scalability Millions of people could use the network for voice, video or ordinary data transmission and the available bandwidth can be divided into several rates. The services are available on a wide range of physical infrastructures;

Reliability Network errors need to be detected and repaired, without the users noticing this, according to the demanded Quality of Service, with a recovery time of 50 milliseconds or less;

Service Management It has to be possible to monitor, diagnose and centrally manage the network with standard, vendor independent implementations and to provide services rapidly. Carrier-class OAM needs to be implemented.

With all background information in mind, the problem definition is stated in the next chapter.

Chapter 3

Problem Definition

In the previous chapter, the background of the main problem is described. From this background, the research question was deducted [2, 5].

This research is based on the following research question:

> *In which way(s) are SURFnet's customers able to setup an end-to-end connection through SURFnet6 and their own internal ethernet network, with use of the new ethernet extensions Provider Backbone Bridges (PBB) and Provider Backbone Transport (PBT)?*

In order to provide an answer to the research question, the functions of the PBB (802.1ah) and PBT (802.1Qay) protocols will be described as well as the operation of these protocols.

3.1 Research Objectives

During this research, the research question, stated in the previous section, will be answered. This will be done by obtaining the following research objectives:

- Create an inventory of the functions and operations of the new ethernet extensions (802.1ah, 802.1Qay);
- Examine how essential properties of lightpath operation can be obtained with the new ethernet extensions;
- Examine possible solutions found to setup a end-to-end connection through SURFnet6 and an ethernet based Local Area Network.

To achieve these objectives, research methods are set-up – see chapter 4.

Chapter 4

Research Methods

In section 3, the research question and research objectives have been described. In order to provide an answer to the research question, research methods are set up and described in this chapter. To get a clear view of the environment of the new ethernet extensions, articles and other documents on these extensions need to be searched for and have to be read.

In the first week of this research, A. Toonk (SARA) and R. van der Pol (SARA) gave a presentation about the current SURFnet6 layout. Also, draft-standards of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) of the protocols 802.1ag, 802.1ah and 802.1Qay are obtained this week. In the second week of this research, interviews with G. Jacobs (Nortel Networks) and P. Grosso (University of Amsterdam) are conducted.

With the articles, IEEE draft-standards and interviews as a source of information, the functions and operation of the IEEE 802.1AB (*Link Layer Discovery Protocol*), 802.1ag (*Connectivity Fault Management*), 802.1ah (*Provider Backbone Bridges*) and 802.1Qay (*Provider Backbone Transport*) protocols are to be described.

The next step is to describe an "ordinary" PBT setup in a simple network environment. With the knowledge of the current SURFnet6 infrastructure, the operation of the new ethernet extensions, and after creating a simple PBT-enabled network, the desired SURFnet6 infrastructure is being examined. Operators of SURFnet6 may assist in answering the question: is it possible to enable PBT in the domains of SURFnet's customers? All the necessary steps for enabling PBT will be described. When PBT is available to customers of SURFnet6, end-to-end connections over ethernet, with carrier ethernet requirements present, can be made.

Chapter 5

SURFnet6

In chapter 2 the background of the research was already described. In this chapter the SURFnet6 network, which is a result of the Gigaport project will be described more thoroughly. It will discuss the topology and techniques that are used at SURFnet6, as well as the reasons for the creation of the network. The purpose of this chapter is to give a clearer picture of the network that is used in this research.

5.1 Introduction

SURFnet6 is the new hybrid SURFnet-network [6]. The network is hybrid, because it supports both traditional IP-routing as well as lightpaths; a lightpath can be interpreted as a deterministic optical point-to-point connection, a circuit-switched path. These lightpaths can transport data up to 10 gigabits per second and provide a certain Quality of Service. SURFnet6 delivers the internet connections for many universities and research institutions. This national network, see picture 5.1, also provides connections to other networks in The Netherlands and abroad. This chapter will describe the most important techniques, protocols and products used by SURFnet6.

Figure 5.1: SURFnet6 Geographically (Source: SURFnet)

5.2 Optical Network

SURFnet6 consists of 6000 kilometers dark fiber, which is as long as the national railroads network in The Netherlands. SURFnet delivers different lightpath speeds to its customers. The customers connect to an (1Gbps or 10Gbps) ethernet interface, to which they can connect their devices. The point-to-point lightpaths that are delivered can be created unprotected, redundant and protected:

- An unprotected path is a single lightpath through the SURFnet6 network;
- Next to the unprotected path the clients of SURFnet can also order a redundant path. The redundant paths are two totally geographically separated paths. The clients whom order these paths have to do their own load balancing and failover when the backup needs to be used;
- A protected path consists of a primary and a backup path. The customers connect to a single interface. In case of a link failure, the network switches to the backup path within 50 milliseconds.

Next to these three forms of point-to-point lightpaths, SURFnet also delivers Optical Private Networks (OPN), for instance to connect multiple locations of a university.

In case of a protected path, SURFnet monitors the paths that are created by their customers, which is done using a control plane. When a problem occurs this control plane will signal it. It is important that a lightpath is monitored because of the high demands – security, reliability and capacity – of the SURFnet customers. It is because of these demands that the topology of SURFnet6 is based on several network rings through The Netherlands; all rings are connected to each other. The topographic layout can be found in figure 5.1.

5.2.1 Lightpaths and SDH

A lightpath through SURFnet6 is created by a point-to-point connection created with Synchronous Digital Hierarchy (SDH). SDH is an implementation of Time-division Multiplexing (TDM). TDM is a technique, which works with timeslots on a line; two or more signals or bit streams are transferred apparently simultaneously as sub-channels in one communication channel. This has the advantage that certain connection speeds can be guaranteed, however, it has the disadvantage that it is not as efficient – when for instance sender A is using its timeslots completely and sender B is not using its timeslots at all; the timeslots for sender B will still be reserved and therefore not used by any sender.

SDH delivers circuit switched connections, which deliver a dedicated circuit between nodes before any communication can occur. Another variant of SDH is SONET, which can be compared to SDH. Generally: SONET is used in North America and SDH is used in the rest of the world.

Table 5.1 shows the protocol stack as used by SURFnet6, it will help explaining the different protocols. All layers in the table will be explained, starting with the lowest layer.

<table>
<tr><td colspan="2">IPv4 or IPv6</td></tr>
<tr><td colspan="2">1GB Ethernet 802.3z or 10GB Ethernet 802.3ae</td></tr>
<tr><td>GFP</td><td rowspan="2">OTN</td></tr>
<tr><td>SDH</td></tr>
<tr><td colspan="2">DWDM</td></tr>
</table>

Table 5.1: Protocol stack of SURFnet6

Dense Wavelength Division Multiplexing (DWDM) is a technique, which is being used for sending and receiving multiple colors, so multiple data streams exist on a single fiber. Different devices are used to separate the different colors in a single fiber (see section 5.2.2). On top of DWDM two different protocols are used, the first is SDH, which is already described in the previous paragraphs, and the second is Optical Transport Network[1] (OTN). OTN is a telecommunication standard that is used by SURFnet6 to put an optical signal on the fiber.

SURFnet6 uses next-generation SDH, this can be described shortly as SDH in combination with Generic Framing Procedure[2] (GFP). GFP allows the signals of higher-layers, for example ethernet. As shown in table 5.1, this is also used by SURFnet6.

[1] OTN is described in the ITU-T standard G.709. More information can be found on the ITU-T website.
[2] Like OTN, GFP is also a ITU-T standard and is described in G.7041

This allows mapping of variable length, higher-layer client signals over a transport network like SDH/SONET. The client signals can be protocol data unit (PDU) oriented (like IP/PPP or Ethernet Media Access Control) or can be block-code oriented (like fiber channel)[7]. The last two layers can be compared to normal ethernet and IP headers.

All of this is obtained with different products. In the next section the most important of these will be examined closer.

5.2.2 Devices

The following equipment is used in the SURFnet6 network:

- Avici Routers for traditional IP-routing;
- Nortel devices for the optical network:

  Nortel ERS8600 used for typical IP connections;

  Nortel OME5200 is being used to connect the ERS8600 to the Nortel CPL. The OME5200 transforms 10GBps Ethernet to OTN based traffic;

  Nortel OME6500 used for typical lightpaths; directly connected to the Nortel CPL;

  Nortel CPL (Common Photonic Layer) is being used for amplifying the DWDM signals and is able to add/drop certain wavelengths of the multicolored lightpaths and for instance severe two lightpaths.

The current SURFnet6 is described in this chapter. To create host-to-host connections through this network with guaranteed Quality of Service, PBT is needed. The next chapter contains a description of the protocols needed for creating a PBT-based environment as in chapter 6.

Chapter 6

PBT Theoretical

To get a better understanding of the Provider Backbone Transport protocol, some of the underlying protocols need to be examined. This chapter will start by describing the 802.1AB protocol – Link Layer Discovery Protocol – which is used by PBT to discover the network layout and forwarding this information to the control plane or management layer. The management layer uses the 802.1ag protocol to monitor the links and trunks in the PBT layout, so this will be described. Because PBT is an expansion of the PBB protocol – 802.1ah –, hence PBB-*TE*, this protocol will be described in depth. Finally the 802.1Qay (Provider Backbone Bridges with Traffic Engineering) or PBT protocol will be described.

6.1 802.1AB, Link Layer Discovery Protocol

On the 28th of March 2005 the IEEE-SA Standards Board approved the 802.1AB standard [8]. This standard is titled "Station and Media Access Control Connectivity Discovery"; the protocol is called "Link Layer Discovery Protocol" (LLDP) and is used in a IEEE 802-based Local Area Network to let the different stations attached discover each other. The purpose of this section is to give a clear introduction to the IEEE 802.1AB protocol. An IEEE 802.1AB enabled device advertises its system information to its neighbours. The neighbors save the system information in a Management Information Base[1] (MIB), which can then be used by a management protocol such as the Simple Network Management Protocol. This way the network advertises its information to a management system, which then knows the network topology, which systems are connected to it, what port status is effective per port, et cetera. The 802.1ad protocol works on all IEEE 802-based LANs, so it can be used on IEEE 802.3 as well as Token Ring. Because it works with standard MIBs the protocol is interoperable between systems from different vendors.

Normally a network management system would discover the network status, so why allow the network to discover itself and then tell it to a management station? The IEEE gives three reasons for the creation of this new standard:

1. Facilitate multi-vendor inter-operability and the use of standard management tools to discover and make available physical topology information for network management;
2. Make it possible for network management to discover certain configuration inconsistencies or malfunctions that can result in impaired communication at higher layers;
3. Provide information to assist network management in making resource changes and/or reconfigurations that correct configuration inconsistencies or malfunctions identified in 2 above.

The next section will describe more about the operation of the protocol, followed by a detailed description of the LLDP-Data Unit (LLDPDU).

[1] The 802.1 standard gives the following information about a MIB (module): *The specification or schema for a database that can be populated with the information required to support a network management information system*

6.1.1 Principles of Operation

An 802-based LAN with LLDP contains agents, which use the services of the Logical Link Control (LLC) and the MAC address to transmit and receive information from and to other LLDP agents. It uses a special multicast MAC address to send the information to other agents. The transmission of the LLDPDU can be triggered by two factors:

- The expiration of a timer, transmit countdown timing counter;
- Status or value changes in one of the information elements in the local system.

When the transmission of the LLDPDU has been started, the local system will put its own – new – information from its MIB in a special packet format called a TLV (Type, Length and Value). The TLVs are inserted into an LLDPDU, which will be transmitted using the special multicast address.

When an LLDP receive module receives the incoming LLDPDU it will recognize the LLC entity and the MAC address. It will then use the information from the different TLVs to fill the LLDP remote systems MIB. The frame format, which is described in the next section, will explain the LLDPDU more thoroughly for a better understanding of the operation of sending and receiving the LLDP frames.

The LLDP agent can transmit information as well as receive information about the remote systems. It is possible to turn either the receiving or the sending functions of the agent off. This makes it possible to configure an implementation, which restricts a local LLDP agent either to transmit or receive only, or to do both.

802.1AB frame format

The frame format of the 802.1AB standard is derived from a normal ethernet frame and contains the following information:

Destination address is the special LLDP multicast address 01-80-C2-00-00-0E;

Source address contains the MAC-address of the sending port;

Ethertype is the LLDP ethertype, 88-CC;

LLDPDU contains different TLVs, which can be found below.

The LLDPDU frame format is somewhat variable, it has some mandatory TLVs, which are always included and several optional TLVs, which can be selected by the network management. The format of the frame is as follows:

1. The following three TLVs are included at the beginning of each LLDPDU;

   Chassis ID TLV identifies the chassis that contains the LAN station associated with the transmitting LLDP agent. It is also used for the MAC Service Access Point (MSAP), which is used to identify a system. Figure 6.1 displays the format of the Chassis ID TLV;

   Port ID TLV contains the port component of the MSAP identifier associated with the transmitting LLDP agent;

   Time To Live TLV contains an integer value in the range 0 to 65535 seconds, which is used to indicate the Time To Live of the LLDPDU. When this time reached zero, the LLDPDU should be thrown away. Normally, the LLDPDU is renewed before this exceeding of time;

2. The optional TLVs can be inserted in any order. The description of each of these different optional TLVs goes beyond the scope of this document. Section 7.3 in [8] describes all optional TLVs;

3. The **End Of LLDPDU TLV** is the last TLV in the LLDPDU, and is used to mark the end of the LLDPDU.

In section 6.4 the use of LLDP in Provider Backbone Transport will be discussed.

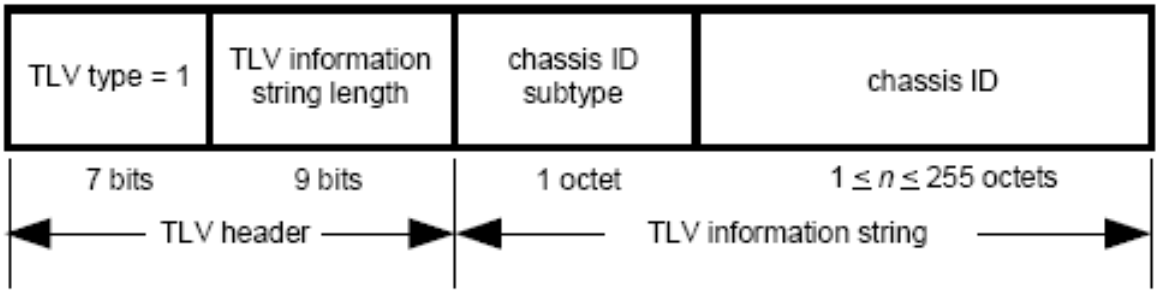

Figure 6.1: Chassis ID TLV Format [8]

6.2 802.1ag, Connectivity Fault Management Protocol

In the past, ethernet was only used as a standard used by Local Area Networks (LANs). Nowadays, ethernet may be deployed in metropolitan-area networks (MANs) and wide-area networks (WANs); see also section 2.1. Ethernet OAM – see section 6.2.1 – is introduced, because the current management protocols do not have the operations required to manage to individual layer 2 ethernet services and do not assist with provisioning of ethernet services, which is required to coordinate the configurations of the respective ethernet equipment [9]. Today's network operators demand implementation of OAM functions, as ethernet shows its capacities to be part of a next-generation infrastructure. The IEEE 802.1ag standard – Connectivity Fault Management – is developed for detecting, isolating and reporting connectivity faults at Provider Bridged networks and Customer-Virtual Local Area Networks. The initial draft for the 802.1ag protocol has been presented to the IEEE in 2004. The most recent version (Draft 8, [10]) is the pre-standard. It is expected that 802.1ag is standardized by the end of 2007.

6.2.1 Operations, administration, and maintenance (OAM)

OAM is a term for protocols, which assist network operators to remotely monitor their networks, by performing management and service operations [11]. By using an OAM protocol, the network and the demands of customers can be managed cost-effectively, because alerts are generated automatically and the location of a fault is known precisely. The OAM protocols are able to automatically send messages containing the status of the network. For example, an operator receives a message when a device went down. When a fatal error occurs on the network, detailed information is sent, allowing the operator to take action immediately. By using the information from the OAM protocol, operators can show their customers all desired information of their connections. A network with few OAM-functions is harder to debug than a network with well-implemented OAM [12].

The International Telecommunication Union - Telecommunication standardization sector (ITU-T) Y.1731 standard for Fault Management and Performance Monitoring, has been completed in May 2006. This standard facilitates not only fault management, but also performance management (by Frame Loss Ratio, Frame Delay and Frame Delay Variation). 802.1ag is different than Y.1731, because Y.1731 has implemented Fault Notification already. This pro-active notification functionality alerts operators on network errors, even before customers are affected [14].

6.2.2 Principles of Operation

802.1ag operates on a single service instance, which has a specific ID (S-ID) that has been added by 802.1ad, as explained in section 6.3. Although 802.1ag operates with a single S-ID, it can handle multiple VLAN-ID (VIDs). This enables the monitoring of specific sections of the network, of which some links are monitored once or more, and other links are not monitored at all.

Fault Management

The IEEE 802.1ag standard is being developed to handle fault management in carrier ethernet networks. This new standard – currently in draft – supports three kinds of fault management for signaling, verifying

and isolating an error in a path [13]:

Fault detection Can be compared to an always-on heartbeat protocol. The fault detection is done by sending a Continuity Check Message (CCM) – figure 6.2. Service failure can be discovered by end-nodes. It continuously checks paths between devices in a carrier ethernet network. The nodes on both sides of the connection are sending CCMs at certain intervals; when a message does not arrive on the other side, a fault is detected.

Fault verification Used to verify the problem exists, after the fault has been detected by *Fault detection*. A LoopBack Message (LBM) is sent to the end-point. The intermediate node replies with a LoopBack Reply (LBR), see figure 6.3;

Fault isolation Used to pinpoint the exact location of the fault, after it is verified, see figure 6.4. The first node sends a LinkTrace Message (LTM) towards the second node on the path, which replies with a LinkTrace Reply (LTR) and sends a LTM towards the third node. This way, the fault can be precisely located.

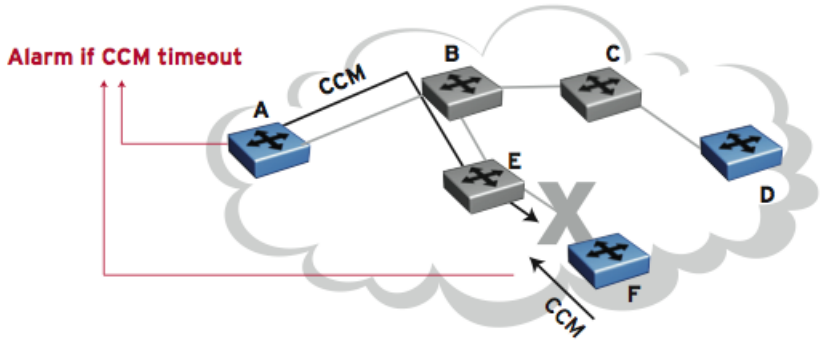

Figure 6.2: Continuity Check Messages - Fault detection in 802.1ag. Source: [13]

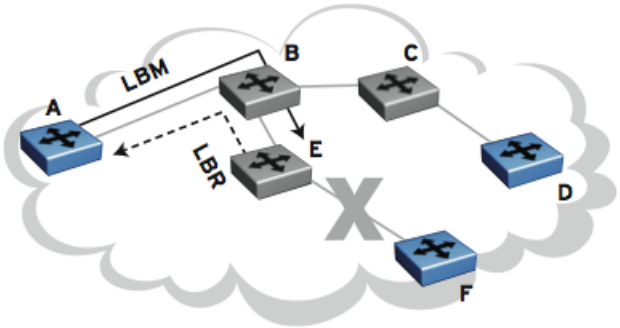

Figure 6.3: Loopback Messages - Fault verification in 802.1ag. Source: [13]

Frame format

The means for identifying Connectivity Fault Management (CFM) Protocol Data Units (PDUs) depend on the medium. For media using a Type/Length field, e.g. IEEE 802.3 media, the identification consists of two octets containing the type value in hexadecimal notation [10].

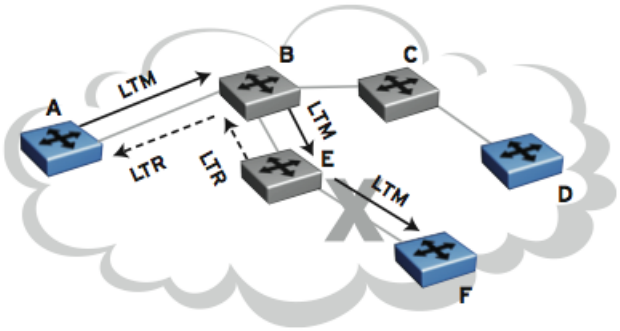

Figure 6.4: Linktrace Messages - Fault isolation in 802.1ag. Source: [13]

The values of these octets are not defined yet as in IEEE P802.1ag Draft version 8 [10].
The common CFM header format is shown in table 6.1.

MD Level (*Octet 1, high-order 3 bits*)
Version (*Octet 1, low-order 5 bits*)
OpCode (*Octet 2*)
Flags (*Octet 3*)
First TLV Offset (*Octet 4*)
Varies with value of OpCode (*Octet 5*)
End TLV (*Octet: First TLV Offset + 5*)

Table 6.1: CFM Header Format [10]

The fields used in table 6.1 are described by the IEEE 802.1ag draft [10]:

MD Level indicates the Maintenance Domain (MD). The field contains a 3-bit integer, which is assigned to the domain. The higher this integer, the more psychical reach the domain has;

Version Protocol version number, with default value 0;

OpCode indicates of which type the PDU is. Possibilities are: Reserved, Continuity Check Message (CCM), LoopBack Reply (LBR), LoopBack Message (LBM), LinkTrace Reply (LTR), LinkTrace Message (LTM) or Reserved by Y.1731;

Flags The Flags field is defined separately for each OpCode;

First TLV Offset The offset, starting from the first octet following the First TLV Offset field, up to the first TLV in the CFM PDU;

Varies with value of OpCode The content of this field is depending on which OpCode is assigned;

End TLV Is a required field, its type is 0.

The content of the TLV (Type, Length, Value) fields is depending on the OpCode. A description of all TLVs can be found in the IEEE 802.1ag draft [10].

6.3 802.1ah, Provider Backbone Bridges

In the IEEE 802.1Q standard [15], Virtual Local Area Network (VLAN) functionality is described. This standard adds the VLAN-ID tag to the ethernet frame header and the ethertype is set to Tag Protocol ID (TPID) 0x8100. With this VLAN tag – also named Q-tag – it becomes possible to divide a LAN into multiple VLANs, with a maximum of 4,094 VLANs. For example, virtual LANs can be used to

split different departments within an organization. The traffic of a certain VLAN is not visible in other VLANs; VLANs introduce hierarchy in a flat network structure. However, the needs for a three-tiered hierarchy – service provider, customer and individual department – were growing.

802.1ah – also known as MAC-in-MAC or Provider Backbone Bridges – completes the future work explained in the 802.1ad protocol (Provider Bridges) standard [16]. This standard uses the ordinary mode of operation of ethernet, so MAC learning, flooding of unknowns, STP, broadcast and multicast are enabled (see figure 6.5). In an article of Nortel Networks the functioning of 802.1ad and the main reason for the development of 802.1ah – further referred as Provider Backbone Bridges (PBB) – is explained:

> "IEEE 802.1ad (also known as Q-in-Q, stacked VLANs or Provider Bridges), extends the original concept of VLANs. IEEE 802.1ad simply adds a new Q-tag that allows the service provider to administer their own tags to identify individual customer networks, while the first (original) Q-tag is used to identify VLANs within the customer's network (i.e. departments in our example). Although Q-in-Q supports a three-tiered hierarchy, the service provider can still only create 4,094 customer VLANs, which is insufficient for large metropolitan and regional networks [17]."

PBB was developed, because a service provider could only create 4,094 customer VLANs. Instead of creating a separate VLAN tag, PBB encapsulates the 802.1ad header in another MAC header, hence MAC-in-MAC. This allows the customer VLANs to be scaled to at least 2^{24} Service Instances [21]. This is due to the fact that the I-SID field, which is 24 bits long, was introduced (the I-SID field will be explained in the frame format section). PBB allows the interconnection of several 802.1ad networks (Provider Bridged networks). This brings more support for large LANs and Metropolitan Area Networks (MAN) [18]. According to Nortel [17], these networks require:

Security – When interconnecting customer and service provider networks the addressing of both sides should be separated for obvious security reasons;

Simpler operations – When one side needs to make changes to its network, this side should not have to worry about the overlapping MAC addresses or VLANs;

Robustness – All tiers are more robust when they are separated (no forwarding loops, broadcast storms);

Lower capital expenditure – MAC addresses should not be traded between the different tiers, because this would cost a lot of memory and processing power.

Currently the 802.1ah standard is in Draft – version 3.4 – and is expected in July this year (2007). The draft already describes the bridge management, for example the SNMP MIB for access to the devices used in a Provider Backbone Bridged Networks (PBBNs). The draft also discusses the Multiple Spanning Tree Protocol (MSTP). The MSTP differs from STP in that it is capable to allow frames assigned to different VLANs to follow separate paths through the network. This is based on different Multiple Spanning Tree Instances, within various regions of LANs. These regions are all connected into a single spanning tree, the Common Spanning Tree (CST). For the operation of PBB the MSTP needed some changes to prevent loops through different PBBNs, without having to link the spanning tree of all PBBNs. The management protocol leans on the 802.1ag protocol, which is described on page 11. This standard also describes the support of the MAC service by PBBNs and the principles of PBBN operation.

6.3.1 Principles of operation

The functioning of the 802.1ah protocol is illustrated with this example: a customer – like in the three-tiered hierarchy – ethernet frame arrives at a Provider Backbone Edge Bridge (PBEB). The PBEB adds a service provider MAC header to the customer ethernet frame. The PBEB checks the service provider MAC address against its forwarding tables and forwards the ethernet frame as if it is an ordinary frame. In the core of the network the switches do not have to be PBB enabled, because these switches see a normal ethernet frame (the one created by the PBEB). The frame travels through the network and

arrives on the other edge at another Edge Bridge. This PBEB will de-encapsulates the extra MAC header, so the customer's frame is again present on the other side of the network. The PBEB forwards the ethernet frame by using the internal forwarding table and the frame arrives at its destination. This is illustrated in figure 6.5. In this figure, the PBEBs are the bridges that connect the customers' networks to the provider network and the customers' networks are also Provider Bridges (802.1ad) enabled.

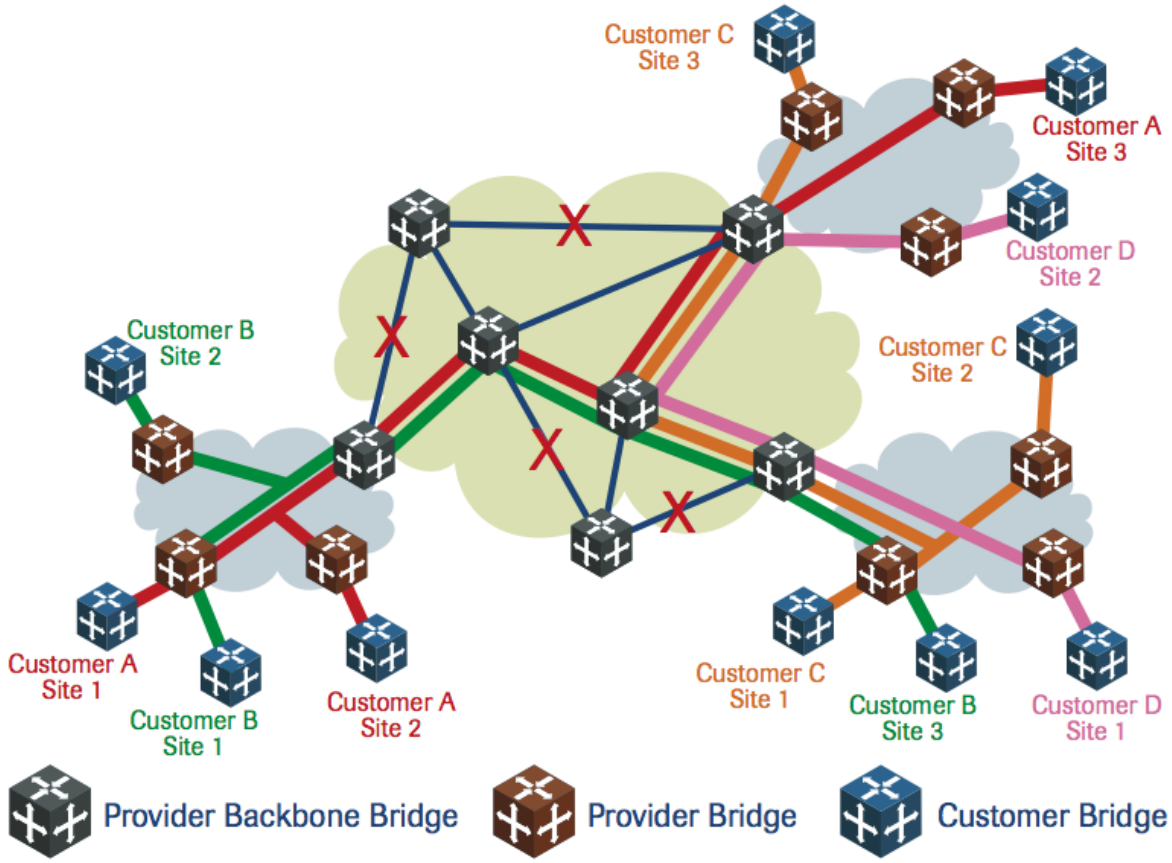

Figure 6.5: Provider Backbone Bridges [19]

This technique has several advantages (as stated by Nortel [18]):

> "This solution allows customer's MAC addresses to overlap with the provider's MAC addresses, because the customers' Service Frames are tunneled by Provider Backbone Bridges and are not used when switching frames inside the provider's network. As a result, customers are free to assign identifier and class of service values to their VLANs, without concern that the service provider will alter them. Meanwhile, the service provider does not need to worry about coordinating VLAN administration with its customers.
>
> Also, because the service provider's switches only use the provider MAC header, there is no need for them to maintain visibility of customers' MAC addresses, reducing the burden on the forwarding tables in the provider's network. This also ensures that changes to the customers' networks do not impact the provider network, improving the stability of the service provider's network. Finally, customer security is improved, because the service provider switches are no longer inspecting the customer MAC header."

The Provider Backbone Bridges tunnels the customers' Service Frames, so control frames, like STP frames, are also tunneled through the provider's network. This allows control protocols, like STP, to be used separately by the customers' networks and the service providers' network. However, as discussed, the STP that is used in the customers' networks, should not interact with the STP used in the service provider's part of the network, which is PBB enabled.

Further operation will be made clear in the next section where the frame format extensions that were added for PBB are described.

802.1ah frame format

Table 6.2 indicates the frame format for 802.1ah [20]. The upper side of the frame (payload to customer-DA) is the part that is used in a Provider Bridges (802.1ad) enabled network. The lower part is the part that is added by a PBEB, when the frame travels through a backbone network.

Payload
Ethertype (802.1Q= 81-00)(*16-bits*)
Customer-VLAN ID (*16-bits*)
Ethertype (802.1ad= 88-a8)(*16-bits*)
Service-VLAN ID (*16-bits*)
Ethertype (standard ethertype, for example ethertype IPv4 (08-00) or IPv6 (86-DD))(*16-bits*)
Customer-SA (MAC)(*48-bits*)
Customer-DA (MAC)(*48-bits*)
I-TAG (*Short I-TAG: 32-bits, Long I-TAG: 128-bits*)
Ethertype (Can be used to specify what is encapsulated)(*16-bits*)
Backbone-VLAN ID (*16-bits*)
Ethertype (specifies whether Long or Short I-TAG)(*16-bits*)
Backbone-SA (MAC)(*48-bits*)
Backbone-DA (MAC)(*48-bits*)

Table 6.2: 802.1ah Provider Backbone Bridges (PBB) frame format

In the table can be seen that the extra MAC header that is added by the 802.1ah protocol can operate as a normal ethernet frame, so it can be used by ethernet devices, which do not have PBB enabled. The backbone source and destination address have the same function as they would have in normal ethernet, just as the Ethertype and Backbone-VLAN ID. The next field contains the instance-tag (I-TAG) [21]. In the previous versions of the protocol, this used to be the I-SID (I-Service ID), but the I-TAG exists of more fields then just the I-SID. This I-TAG comes in two flavors, which are used by the PBEBs, the Short Service Instance TAG and the Long Service Instance TAG:

- Short Service Instance TAG: The short TAG exists of the 24-bit I-SID, which identifies the service in the provider's network, and 4 control fields:
  - Priority, is used for the customer priority;
  - Drop_eligible is a 1-bit field that indicates the customer drop eligibility;
  - Res1 is a 2-bit field that is used for any future format variations;
  - Res2 is also a 2-bit field that is reserved for future format variations;
- Long Service Instance TAG consists of the Short Tag, including the customer source and destination address.

The purpose of the Long Service Instance TAG is to indicate that an ethernet frame is encapsulated inside the PBB header, while the Short Service Instance TAG is intended for multi-protocol encapsulation. The fields are distinguished by a different ethertype. At the moment of writing this report, the values of the ethertypes still have to be assigned[2]. The Ethertype (before the I-TAG field), can be used by PBEBs to known what kind of frame is encapsulated (DIX, 802.3, etc).

Based on a figure found in a Nortel Networks document [17], the overview found in figure 6.6 was created. It describes the transition of the 802.1Q header format to the 802.1ah header format and can be used as a summary of the previously explained protocols and header formats. In the next section PBT will be explained, where the 802.1ah header is also used. This figure can be used as a reference.

[2]Based on PBB draft 3.4, March 2007

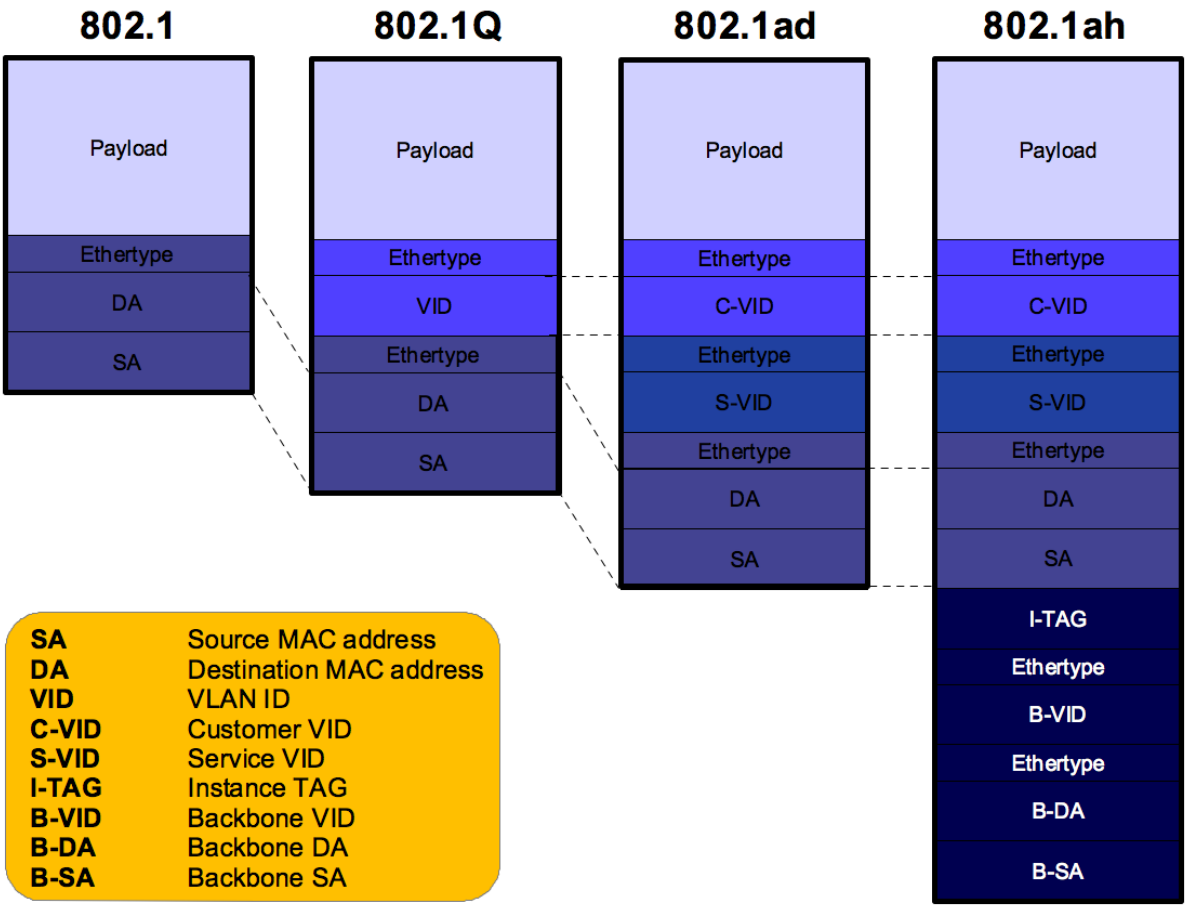

Figure 6.6: Overview of explained headers

6.4 802.1Qay, Provider Backbone Transport

Provider Backbone Transport (PBT) is a combination of ethernet extensions, created to establish a carrier ethernet, see page 2. PBT – named 802.1Qay by the IEEE – is based on the extensions 802.1Q, 802.1AB, 802.1ad, 802.1ah and 802.1ag. The relation between these protocols is illustrated in figure 6.7. During this research, the IEEE 802.1Qay Draft 0.0 [22] was obtained.

The use of standard ethernet in a WAN or MAN is not possible without a modification or extension of the protocols being used. This is due to the fact that ethernet is a connectionless protocol. Connectionless means that ethernet is designed to transport ethernet frames toward their destination, somewhere in the network. When the destination is unknown, a layer two device will send the data on all of its outgoing ports (called flooding of unknowns). As the destination is reached through one of the ports, it will reply; the switch now knows this destination address can be found behind that particular port and stores this information in its memory (MAC learning). When using flooding in a large network, it is clear congestion occurs immediately.

In order to use ethernet in a MAN, Nortel Networks [17] states that service providers need a way to:

- Deliver guaranteed, deterministic services over Ethernet infrastructure on a wider scale;
- Ensure reliability, management and scalability in order to deliver the multimedia services that enterprises demand;
- Take advantage of the operational and cost efficiencies that Metro Ethernet can generate while moving towards a converged infrastructure.

PBT is being developed with scalability and offering Quality of Service in mind. In [17], the need for PBT is described:

> "In todays MAN, conventional wisdom has been that connectionless Ethernet needs a 'helper', just as IP did. So it is only natural that many have looked to solve the problem by extending MPLS into the MAN. This strategy worked well for the relatively few number of nodes (i.e. hundreds) in the WAN — but it quickly becomes unmanageable and expensive for the relatively large number of nodes (i.e. thousands) in the MAN. The problem is that deploying MPLS requires implementing many new protocols and standards (...) which adds not only operational complexity and costs, but also increased network equipment capital costs due to the likely control plane and data plane upgrades required to support them. Service providers are seeking to migrate their networks onto a purely packet infrastructure and their goal is to combine the flexibility of packet processing with the determinism, OAM and operational attributes which they are used to from circuit infrastructure — all at Ethernets cost points. Service providers have had to deploy 'MPLS everywhere' as there has been no other acceptable or practical alternative for their metro networks — until now."

To use ethernet as a transport technique in a MAN, it needs to have OAM functionality – see section 6.2.1. PBT is able to use 802.1ag, 802.3ah[3], 802.1AB, G.8031[4], Y.1731 (Ethernet OAM) and Metro Ethernet Forum (MEF) Ethernet Performance Monitoring.

When using PBT, it becomes possible to set up a path with QoS and a backup path over an ethernet-based network. The requirements for such a path/tunnel were already described in section 2.1.1. When PBT tunnels are created, appropriate bandwidth is reserved, which supports the provisioned QoS metrics that guarantee SLAs will be met without having to over-provision network capacity as outlined by Nortel [17]. This provisioning needs to be handled by the management layer of the network. The precise implementation of the provisioning is not yet described in the IEEE 802.1Qay Draft version 0.0 [22]. The segmentation of the network is done by using VLANs. A single physical part of the network can have PBT, PBB and normal ethernet frames on different VLANs.

[3]Ethernet in the first mile (`http://www.ieee802.org/3/efm/`

[4]ITU-T Ethernet protection [23]

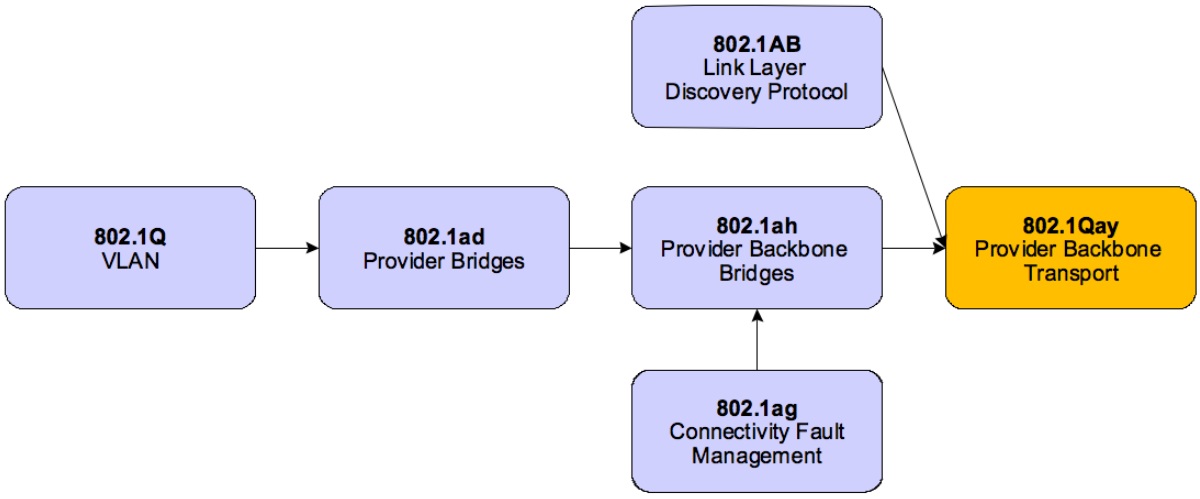

Figure 6.7: Overview of protocols used by PBT

6.4.1 Principles of Operation

With PBT, flooding of unknowns, broadcast, multicast and MAC learning functions are disabled. PBT also ignores spanning tree states, however STP can be used on the non-PBT-VLANs. The network management is responsible for filling the forwarding tables of the layer 2 devices. The management does not determine what the network layout is like; instead, 802.1AB tells the management how the network layout looks like. This makes provisioning and adding nodes much simpler. The forwarding of the packets is based on explicit routes through the network. When the Continuity Check Messages of the 802.1ag protocol are lost in the network, a fault has occurred. In this case, the VLAN-ID (VID) at the sending host will be changed to the backup VID. This process takes place in around 50 milliseconds [24].

802.1Qay frame format

The frame format of PBT is exactly as the format used for implementing PBB, see table 6.2. The difference is in the meaning of the frame's fields. The VID and the B-DA fields together form a 60-bits globally unique identifier. PBT packets will be switched on VID and destination MAC address. There are two VLANs when path protection is enabled: one VLAN provides the current work path and the other provides the backup path. The control plane is used to manage the forwarding tables of the switches; to create PBT tunnels, all switches need to be controlled from one (PBT) domain. This is necessary for the control plane to fill the forwarding tables of the switches and thus set up a path. Because the control plane fills the forwarding tables, learning, flooding and spanning tree have become obsolete for PBT. This technique enables circuit-switching on an ethernet. There is a total of 4,094 VLANs, which can be set up according to the 802.1Q standard. When using for example 64 VIDs for PBT, the other 4030 can be used for "ordinary" connectionless ethernet.

In this chapter, the theory behind PBT and its related protocols has been discussed. Chapter 7 will describe three possible solutions for implementing PBT in a network.

Chapter 7

PBT Practical

The previous chapter described the theoretical background of PBB and PBT. In order to give a complete overview of PBT, three practical environments are described in this chapter. It will start with an example of an ordinary PBT implementation in the provider's domain (figure 7.1). Next will be described how it is possible to use a PBT solution where the customer's network also has PBT enabled to create a carrier ethernet, which includes the customers' networks. Afterwards a situation with two customer PBT enabled domains, inside the customers network, with a tunneled provider network. The application of this last situation to SURFnet6 will be described in the next chapter.

7.1 Implementing PBT

This section is an explanation of figure 7.1. Using this figure, an ordinary PBT setup will be described. The left side of the figure indicates customer A. This customer has a network with a normal ethernet LAN and is connected to the provider. This connection is a Provider Backbone Edge Bridge – PBEB *A*. The provider network contains several other Provider Backbone Core Bridges –PBCBs *C* and *D*. Customer B is connected to the provider network through PBEB *B*.

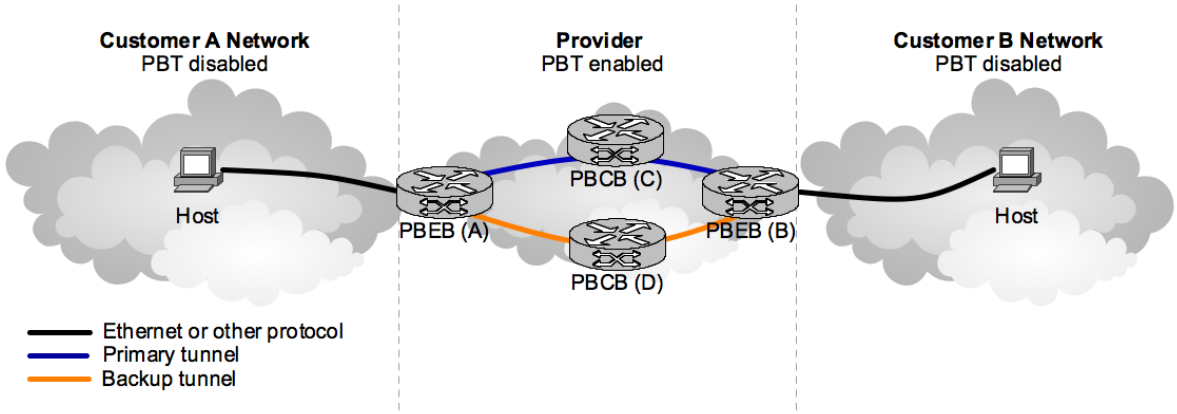

Figure 7.1: Provider PBT solution

A management layer or a suitable control plane manages the core network[1] with the PBEBs and PBCBs – see chapter 5. This control plane fills the forwarding database in the PBCBs with the right entries, so they are able to forward traffic through the path/tunnel. For example, PBCB *C* has an entry for PBEB *B* in combination with a certain Backbone VLAN-ID – the 60-bit identifier – to forward traffic on a certain port in the direction of PBEB *D*. These tunnels are both created by the management layer; this allows a path to be created manually by an operator through this management layer, bandwidth to

[1] As illustrated in figure 7.1.

be reserved, and requirements for a service to be set. The 802.1Qay Draft version 0.0 does not contain any information about the management layer at this moment.

Inside the providers' network, the management layer has created tunnels. When a customer from site A sends traffic to a customer at site B, its destination will be reached as following:

1. The traffic originating from customer A is sent through the customers network to the PBEB *A*. This traffic has no Quality of Service, as the customer's network is based on ordinary ethernet;
2. PBEB *A* has been configured to add the 24-bit I-SID of a certain value to the I-TAG field, based on the S-VID (802.1ad) of the customer's frame;
3. With the use of this value a tunnel – and its backup tunnel – can be selected. These tunnels have to be reserved before the tunnel becomes operational. The tunnels are identified by the 60-bit identifier;
4. The PBEB also adds the backbone destination address of PBEB *B* and its own backbone source address and the backbone VLAN-ID, also based on the S-VID of the customer's frame;
5. The encapsulated frame will be sent through the tunnel and will arrive at PBCB *C*;
6. PBCB *C* has a forwarding entry – which it selects based on the B-VID – and will forward the frame to PBEB *B*;
7. PBEB *B* will de-encapsulate the frame and forward the frame based on the I-SID, which is associated with a certain VLAN-ID;
8. The frame will be delivered to customer *B*.

Section 6.2 already explained that the paths are being monitored by the 802.1ag protocol, so when an error occurs the backup tunnel will automatically be used – with a flip-over time of 50ms. It is obvious that a backup path has to be reserved before operation, to guarantee this failover.

It is possible that inside the customers' network another protocol like Provider Bridges (802.1ad) is enabled [19, 25, 26], however for the combination of PBT with SURFnet6, it is required that the PBT protocol is enabled in the customers' network.

7.2 A larger carrier ethernet

In the previous section a PBT solution inside the providers' network was described. For implementing a complete host-to-host connection with QoS, failover, scalability and service management, another solution has to be examined. This is possible by implementing PBT inside the customer's network. This solution is illustrated in figure 7.2.

This solution delivers the demands for carrier ethernet to the complete connection from host to host, however, it also has several disadvantages. Because it has multiple PBT domains, three control planes exist, in order to ensure the PBT tunnels to be created and ended correctly. Also, every time a frame travels from one domain to another the frame has to be de-encapsulated and encapsulated. In the figure can be seen there are two additional points of (de-)encapsulation – compared to figure 7.1, which gives some overhead. In figure 7.2 the PBEBs create single points of failure in the connections. This could be solved by installing additional PBEBs between the domains. In this situation, there are three domains, monitored by 802.1ag (Connectivity Fault Management) per domain. When a link error occurs on one domain, the other domains will not know the problem has occurred. To solve this issue, the management layer should be handled on another (OSI-) layer. Another host-to-host solution is illustrated in figure 7.3. This single-domain solution also has its specific advantages and disadvantages. For instance, one of the advantages is that only one control plane is required, but a disadvantage is that this solution has less scalability for the different partners. Table 7.1 indicates the differences between both solutions.

As stated before, the encapsulation is a disadvantage of the multiple-domain solution. In a single-domain solution only one control plane or management layer can be used to create the paths through the different networks. However, the different partners have more flexibility when they use their own control plane, like in the multiple-domain solution. This also has an impact on the addresses that are

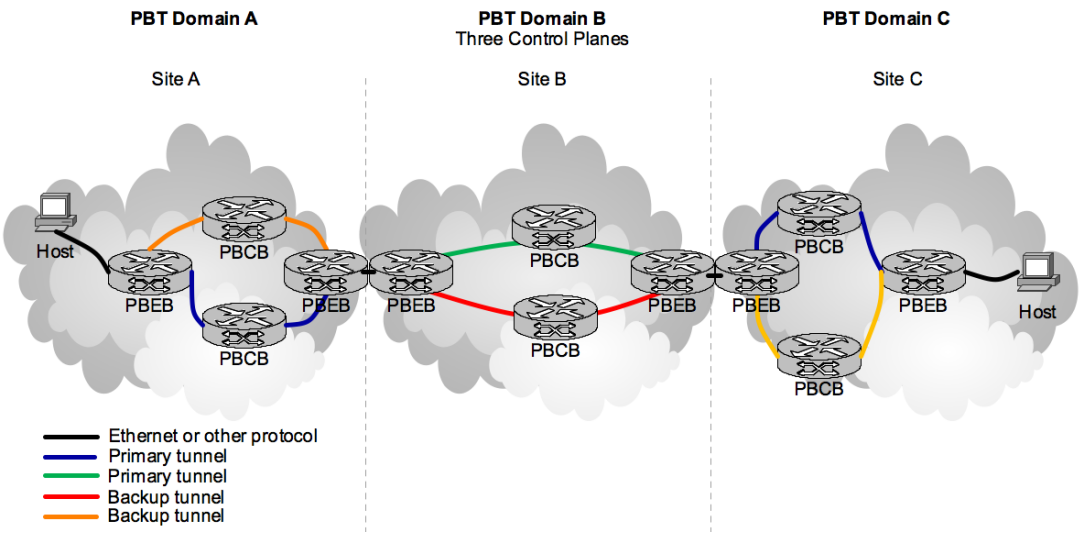

Figure 7.2: PBT solution including the customer's network

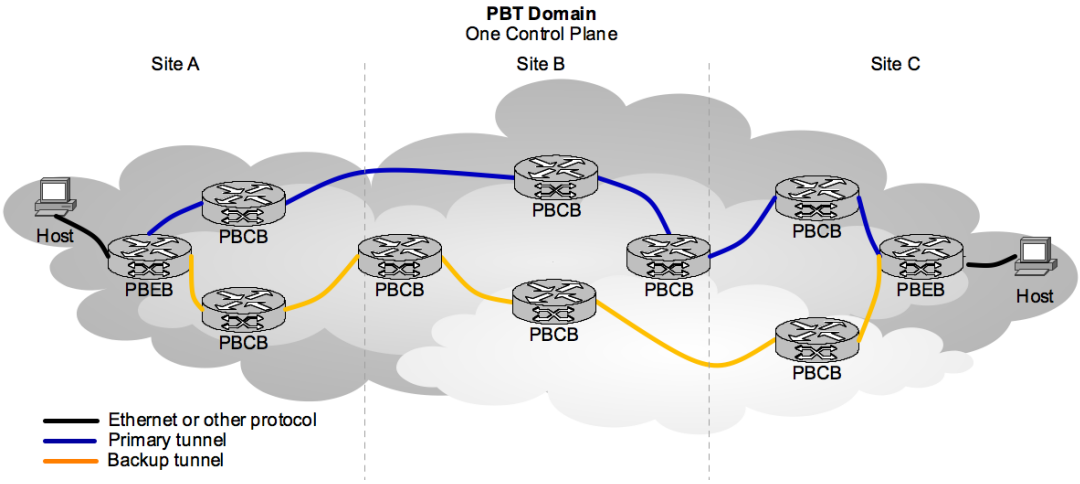

Figure 7.3: PBT solution including the customer's network

used inside the networks. All customers should use the same I-SIDs and 60-bit tunnel identifiers, so agreements have to be made on the use of the fields and identifiers. Because of these requirements in a single-domain solution, this solution has a lower scalability. It is harder to add PBT domains, when all these agreements have to be made with every new customer. This is not so limitless as with the multiple-domain solution, when the requirements of a carrier ethernet must be obtained.

Both solutions have their advantages and disadvantages and both deliver host-to-host carrier grade ethernet. It depends on the situation which solution is the best for a certain implementation. When a provider has a lot of influence in the customer networks a single-domain could be preferred, however, it is possible that a customer will want to operate independent of the provider. In the next section a solution without a PBT enabled core will be discussed, where the customers have PBT enabled and the provider network works with different protocols and standards than the customer networks.

	Single-Domain	Multiple-Domain
(De-)Encapsulation	once per tunnel	once per domain
Management	just one control plane	multiple control planes
(Backup-)paths	many and easier	fewer and harder
Addresses	agreed in consultation	own solution
Scalability	lower	high
Autonomity	dependent of other partners	independent/autonomous
Redundancy	backup paths	single points of failure

Table 7.1: Comparison of two host-to-host solutions

7.3 Situation without a PBT enabled core

Figure 7.4 indicates a PBT implementation where the core network does not have PBT enabled. This core network has its own protocols and standards. Customers, who have PBT enabled, are both connected to this provider network. This can be compared to the multi-domain solution in the previous section, however, this situation has different services; the core of the network has its own implementation of protocols and standard to deliver the specific requirements of a carrier grade ethernet.

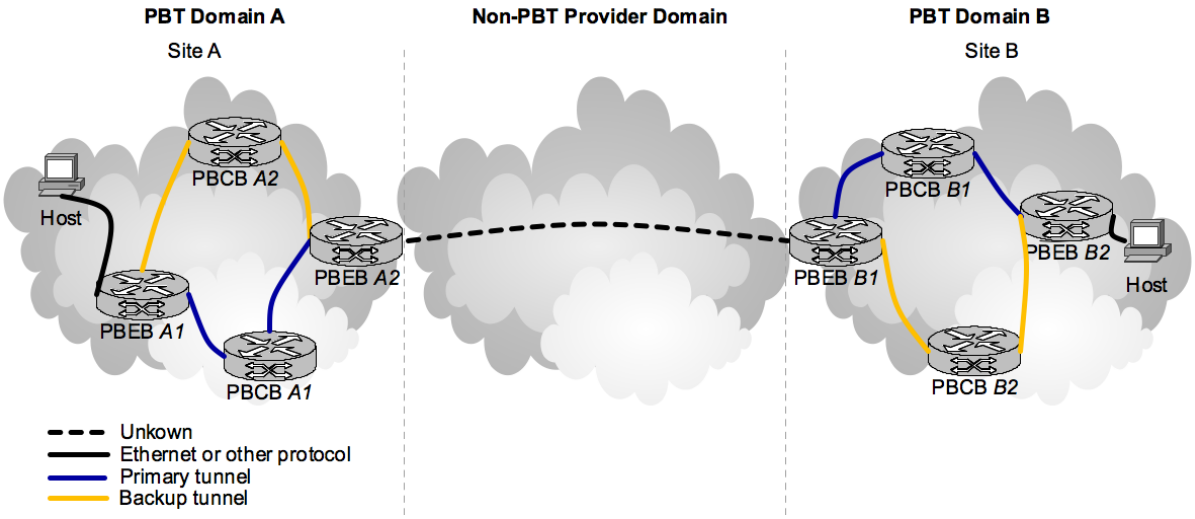

Figure 7.4: PBT solution including *only* the customer's network

Just like the first example of a PBT implementation, an example of PBT traffic-flow will be used, to explain the differences between this implementation and the previous. When frames travel through a network with this last solution, several differences with a normal PBT solution can be seen. The following enumeration indicates a frame that travels from customer A to customer B.

1. The traffic originating from host A is sent through the customers network to the PBEB *A1*;
2. PBEB *A1* has been configured to add the 24-bit I-SID of a certain value to the I-TAG field, based on the Service-VLAN ID (802.1ad) of the customer frame;
3. With the use of this value a primary tunnel and a backup tunnel can be selected – these have to be reserved before. The tunnels are identified by the 60-bit identifier;
4. The PBEB also adds the backbone destination address of PBEB *A2* and its own backbone source address and the backbone VLAN-ID, also based on the S-VID of the customer frame;
5. The encapsulated frame will be send through the tunnel and will arrive at PBCB *A1*;

6. PBCB *A1* has a forwarding entry, which it selects based on the B-VID, and will forward the frame to PBEB *A2*;
7. PBEB *A2* will de-encapsulate the frame and forward the frame based on the I-SID, which is associated with a certain VLAN-ID, so an *ordinary* 802.1Q frame remains;
8. The frame will be encapsulated in other headers – for example SDH – and transported through the network;
9. PBEB *B1* will get the frame and identify it based on its VLAN-ID;
10. The PBEB encapsulates the frame and sends it through the PBCBs and PBEB at the network of customer B;
11. The last PBEB will de-encapsulate the frame and deliver it at its destination;
12. The frame will be delivered to customer B.

Of course, both PBT paths/tunnels should be created before data can travel through these tunnels with the specific requirements. The management layers of the customers can create these tunnels, but it may be possible the customers combine their management layers and create a coordinating solution that manages both PBT tunnels – this is described in section 9.1. It could also be possible that the provider also joins this coordinating solution and that a single-domain control plane is created.

In the next chapter a – possible – solution for SURFnet6 in combination with PBT will be discussed; to deliver host-to-host connection, through a lightpath based core network, in combination with PBT customer domains.

Chapter 8

SURFnet6 and PBT

To establish a host-to-host connection – with reliability, service management and Quality of Service – through SURFnet6 (see chapter 5), PBT could be implemented as in the last solution, mentioned in the previous chapter. This solution has two different PBT domains on both sides of the provider network. When SURFnet would be allowed by its customers to make changes on both sides of its network and thereby implements its demands in the networks of the customers, it would be possible to create a single-domain solution with only one management layer. However, in case of SURFnet it is not possible to have any influence inside the networks of its customers. This report will only discuss the solution with multiple-domains, because SURFnet or SARA is not allowed to make changes in their customers' networks. Hence, the multiple-domain solution for using PBT is the most applicable to the SURFnet6 network.

Figure 8.1 illustrates a solution, which could be implemented with the existing SURFnet6 infrastructure in mind. The figure shows the SURFnet6 core network in the middle and two customer networks on the left and right side. At this moment, the customers' networks are based on ordinary ethernet, without the new extensions mentioned in chapter 6, which can never meet the requirements stated in section 2.1.1. The two grey optical devices deliver ordinary ethernet frames for the connection between the customers. The blue PBT bridges do not exist in the current situation, however, they may exist, but do not support PBT yet.

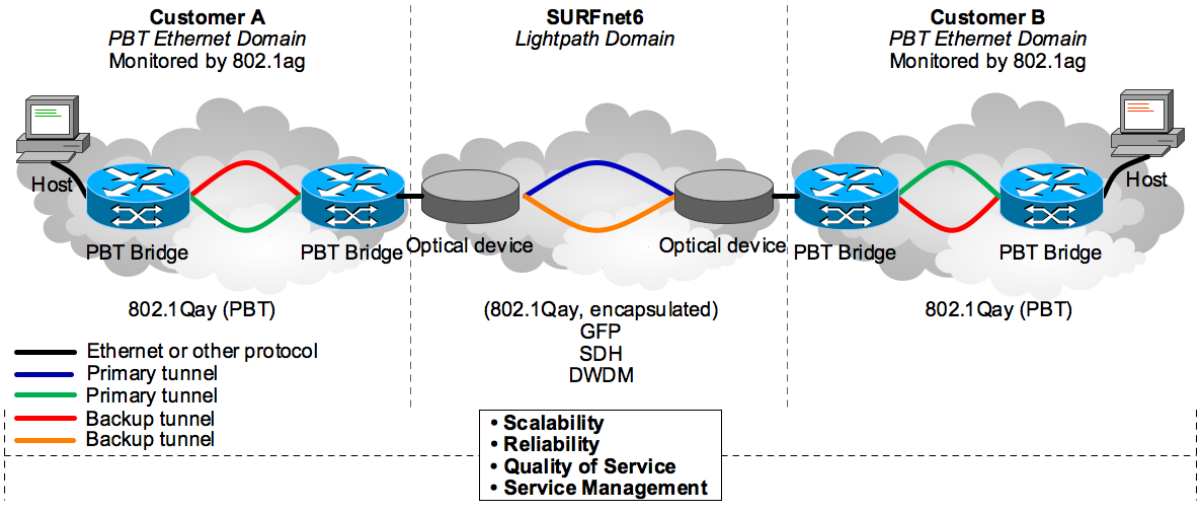

Figure 8.1: Overview of PBT in combination with SURFnet6

To setup a complete host-to-host path that meets the requirements, PBT is implemented in the customers' networks: upgrading (or placing) two or more PBT bridges in each customer's network can do this. Figure 8.1 represents these by the blue bridges. Inside the customers' networks, PBT is used for setting up primary and backup paths.

So in this solution, a Provider Backbone Edge Bridge (PBEB) has to be enabled in the customer's domain/network. This PBEB will terminate the customer's PBT tunnels and have a dedicated connection to SURFnet's optical devices (Nortel OME6500 or the Nortel ERS8600 in combination with the OME5200 as described in chapter 5). This has the advantage that the management planes are completely separated per domain, and the disadvantage that the PBT devices have to be installed, raising the costs per customer. However, on some existing devices, new firmware can be loaded to enable PBT, eliminating the extra costs.

8.1 Implementation

At this moment SURFnet delivers a 1-gigabit or 10-gigabit connection to a customer through its optical devices. These connections deliver the 1 and 10 gigabit links with the use of 802.3ae and 802.3z standards. The optical devices deliver normal ethernet to the customers' PBEBs. The customers need to have two or more PBT enabled bridges to make a complete PBT tunnel inside their network. This can either be through a connection between two PBEBs or by traveling a path through multiple devices (PBEBs or PBCBs). All devices between the host and SURFnet's optical devices need to support PBT – figure 8.1 illustrates this with the blue bridges. Nevertheless, the customer can still use its normal – *old* – hardware for non-PBT LAN operations on the other VLANs.

The following steps describe the implementation phases the customers should take to interconnect:

1. A static lightpath between the two customers already exists;
2. The customer needs to install the PBT bridges in its network;
3. On all the customer's PBT devices, the correct manufacturer software for enabling PBT has to be loaded;
4. When 802.1ad is used inside the customer's network, VIDs and SIDs need to be configured;
5. The management layer (containing at least 802.1AB and 802.1ag) has to be set up;
6. The 802.1AB protocol provides the management layer with information of the network layout;
7. The network operator manually sets a path/tunnel through the customer's network, with an assigned VID and/or SID. This path is terminated at the last PBEB before SURFnet's optical devices;
8. The other customer takes the same steps as described;
9. The link can be activated and a host-to-host connection with Quality of Service is present.

The lightpath through SURFnet6 is created statically in the previous enumeration; it has been configured once, and is not changed afterwards. At this moment, most lightpaths through SURFnet6 are static. However, the lightpaths could be created dynamically in the future, for instance with use of DRAC. Creating dynamically lightpaths increases the complexity of the host-to-host connection. When the connection has to become active, the lightpath should already be reserved and set up. It is possible that no valid circuit can be calculated through SURFnet6. When this occurs, no host-to-host connection can be guaranteed. This may be caused by another customer which reserved a needed lightpath.

8.2 Traffic flow

When a customer enables its network for PBT, two or more PBT bridges have to be installed. In figure 8.1, only two bridges are placed per customer. When expanding the customer's network with more PBT bridges, a situation as in figure 8.2 exists. This is a simplified illustration of the reality. In a normal situation the primary and backup path will take a separated route, so they will offer reliability. The traffic flow through this expanded network is explained below.

The traffic through the created paths follows the route as illustrated in figure 8.2:

1. Host *A* sends an 802.3 ethernet frame to PBEB *A1*, which is connected to PBCB *A1*;

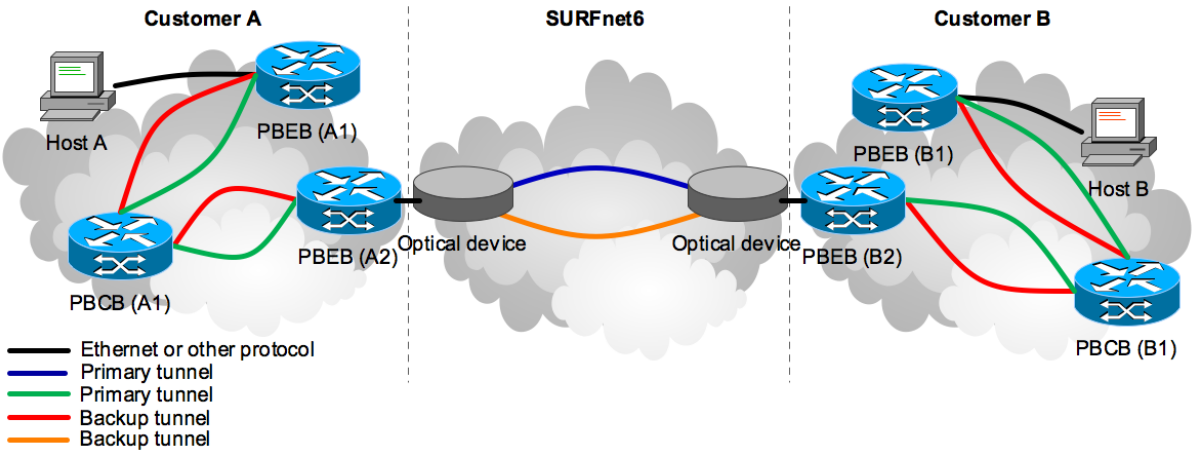

Figure 8.2: Customers' traffic through an extended customer LAN and SURFnet6 (Simplified)

2. PBEB *A1* – to which the host is connected – receives the frame and puts it inside a reserved PBT tunnel towards PBCB *A1*;
3. PBCB *A1* knows where to forward the frame to and forwards it towards PBEB *A2* through the PBT tunnel;
4. PBEB *A2* de-encapsulates the frame and passes it through the SURFnet6 lightpath to PBEB *B2*;
5. PBEB *B2* receives the frame, encapsulates it and puts it inside the PBT tunnel towards PBCB *B1*, with a reserved VLAN-ID and an I-TAG;
6. PBCB *B1* receives the frame and puts it inside a reserved PBT tunnel towards PBEB *B1*, to which Host *B* is connected;
7. PBEB *B1* de-encapsulates the frame. The payload is delivered to host B with the assigned Quality of Service.

8.3 Control Plane

In the PBT solution as illustrated in figure 8.2, there are three management layers: one of customer A, one of SURFnet6 and one of customer B. When a customer wants to create a path, this has to be done in three steps, which is not very efficient. Therefore a solution could be sought where all three management layers, which create the paths, are combined. Customers and provider (SURFnet) need to make arrangements, because the overruling control plane has management capabilities in all networks. One of the possibilities is combining the management layers of PBT with the current SURFnet6 management layer. Another development, focusing on this uniform control plane is named Provider Network Controller (PNC). Soapstone Networks, the software developer behind this control framework, is a business unit of Avici Systems and is focusing on the idea of separating the network control from the network layer, thus separating the control plane.

This chapter discussed a new idea for implementing PBT in the SURFnet6 environment based on the theory as in chapter 6 and the possible solutions from chapter 7. With this information, the entire research question of this research can be answered, which will be done in the following chapter.

Chapter 9

Conclusions

When searching the internet for PBT, a lot of articles can be found about this new protocol standard. This indicates the demand for bringing Quality of Service to ethernet; therefore enabling the technique, which is a result of combining the 802.1Q, 802.1ad, 802.1ah, 802.1ag and 802.1AB protocol standards, to be used in Metropolitan Area Networks. These protocol drafts and standards are described in chapter 6, to explain the theoretical background of the Provider Backbone Bridges (PBB) and Provider Backbone Bridges with Traffic Engineering (PBT). This chapter contains an answer to the research question stated in chapter 3:

> *"In which way(s) are SURFnet's customers able to setup an end-to-end connection through SURFnet6 and their own internal ethernet network, with use of the new ethernet extensions Provider Backbone Bridges (PBB) and Provider Backbone Transport (PBT)?"*

In this report, the theory behind the 802.1Qay (Provider Backbone Transport) protocol is described in chapter 6. Because the IEEE 802.1Qay protocol is currently in draft version 0.0, the exact mode of operation is still unclear. However, several organizations are already testing implementations of PBT on special deployed network sections.

Section 2.1.1 stated all requirements needed to deploy a carrier ethernet with guarantees of service in mind. Using PBT, offers the opportunity to implement these requirements (Quality of Service, Scalability, Service Management, Standardized Services, Reliability) into an ethernet-based network.

Earlier, this document discussed practical implementations of PBT in one or multiple domains, in chapter 7. With the solutions mentioned in this chapter, it becomes clear it is possible to create a PBT based host-to-host connection through a provider network, which is either based on ethernet or not. These solutions deliver performance, reliability and scalability like the lightpaths of SURFnet6 do.

In chapter 8 it is described how PBT can be implemented with SURFnet6, which is described in chapter 5, without introducing impossible changes to the network. A solution to implement PBT into networks of customers of SURFnet is to upgrade (or place additional) LAN equipment, at least two PBT Bridges (PBEBs, see figure 8.1). With this solution, the existing infrastructure can remain intact – or the new devices operate simultaneously with the standard equipment. The result of implementing PBT into the customers' networks is illustrated in figure 8.2.

By enabling PBT on LAN equipment into SURFnet's customers' networks, it becomes possible to create end-to-end connections, thus combining the functionality of PBT and optical lightpaths, and therefore deliver guaranteed quality of service, reliability and scalability over the entire connection with three separate control planes.

Section 7.2 explained the PBT tunnel has to be managed per domain from PBEB to PBEB. In order to do this, the customer has to install two PBEBs between SURFnet6 and the customer's host.

During this research, it became clear that PBT is not finished and tested enough to be implemented in a real world network yet. Some improvements by this new protocol and recommendations to implement these are described in the next section.

Altogether, SURFnet6 could use the new developments of PBT to extend the functionality of their optical lightpath based network, to deliver host-to-host connections, which meet the requirements of carrier ethernet and create a Metropolitan Area Network.

9.1 Recommendations and future work

In the solution described in chapter 8, three control planes co-exist. In order to provide easier management over the complete end-to-end link, it is recommended that these planes are merged into one. New developments like Soapstone Networks' *Provider Network Controller* [28, 29, 30] are aiming to merge the management layers together.

When these layers are merged into one, a single PBT domain, with a transparent provider part, can be created. This solution is illustrated in figure 9.1. The PBEBs, that used to de-encapsulate and end the PBT tunnels, are replaced by Core Bridges. SURFnet's lightpaths form a transparent network, consisting of virtual fibers. In this solution, the PBT frames will be encapsulated in – for example – SDH frames; instead of de-encapsulating the frames. The (dis)advantages of this solution compared to the multiple domain situation, can be found in table 7.1.

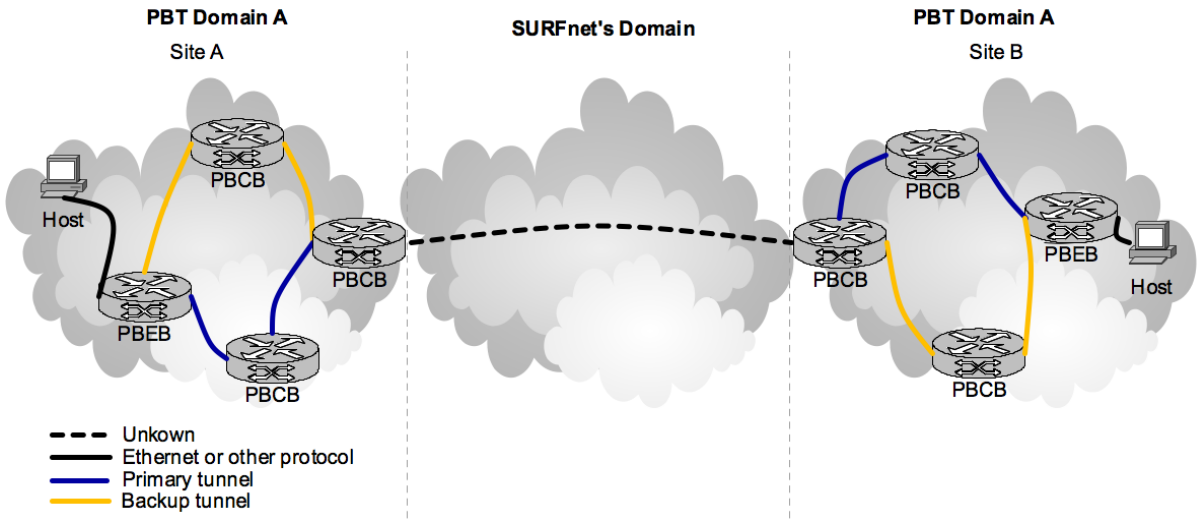

Figure 9.1: Connecting two Provider Backbone Core Bridges

Another improvement could be reached by implementing PBT (-enabled ethernet) in SURFnet6 infrastructure itself. P. Grosso stated that the University of Amsterdam will start a research that will focus on the extension of PBT into campus networks, same as described in sections 7.2 and 7.3. The most likely core network would be SURFnet6, however, this is not necessary. This project, which will take one year, will be started in September 2007. A possible outcome of this research could be that when PBT is implemented in the core of SURFnet6, it is easier to create more reliable host-to-host connections. This could offer a single-domain solution, as described in section 7.2.

Other developments, like Transport MultiProtocol Label Switching (T-MPLS) [31, 32, 33], also offer carrier ethernet-like functionalities. This research is based entirely on describing PBT and its possible solutions. Further research could be conducted in order to find and describe the differences and possibilities between implementing T-MPLS and PBT.

A recommendation for SURFnet when implementing PBT, is to migrate in several small phases: (1) SURFnet can use this report to analyze whether PBT is interesting for their customers. (2) When this is interesting, customers will be advised to look into PBT. (3) The customers can then decide themselves whether they want to use PBT. (4) After this phase, customers contact their vendors for PBT support on their existing equipment. When new equipment is bought by customers, they should try to buy equipment that has PBT support. (5) After the equipment is upgraded, the customers gradually introduce PBT.

This cooperation between SURFnet and its customers will evolve into a large network, in which all customers are able to set up reliable and scalable end-to-end connections, which offer complete Quality of Service.

Bibliography

[1] SURFnet, January 2007, *Toepassingen van dynamische lichtpaden in SURFnet6*[DUTCH], http://netwerk.surfnet.nl/info/dynamische_lichtpaden/toepassingen.jsp

[2] G. Jacobs, May 2007, *Can Ethernet Become the Lingua Franca of the Future Transport Infrastructure?*, http://tnc2007.terena.org/programme/sessions/show.php?sess_id=70

[3] P. Bottorff, April 2006, *Highly Scalable Ethernets*, http://www.itu.int/ITU-T/worksem/ngn/200604/presentation/s7_bottorff.pdf

[4] Metro Ethernet Forum, May 2007, *Carrier Ethernet - 5 Attributes*, http://www.etherwiki.org/Carrier_Ethernet/Carrier_Ethernet_-_5_Attributes

[5] R. van der Pol, 2007, *Research Description*, http://staff.science.uva.nl/~delaat/sne-2006-2007/index.html

[6] SURFnet , January 2006, *Lichtpaden van SURFnet6 markeren nieuw tijdperk voor het internet*[DUTCH], http://www.surfnet.nl/info/artikel_content.jsp?objectnumber=107219

[7] Wikipedia, April 2007, *Generic Framing Procedure*, http://en.wikipedia.org/wiki/Generic_Framing_Procedure

[8] IEEE Computer Society, May 2005, *Station and Media Access Control Connectivity Discovery*, http://standards.ieee.org/getieee802/download/802.1AB-2005.pdf

[9] Cisco Systems, Inc., September 2006, *Ethernet Operations, Administration and Maintenance*,http://www.cisco.com/en/US/products/hw/routers/ps368/products_white_paper0900aecd804a0266.shtml

[10] IEEE 802.1 Working Group, February 2007, *IEEE P802.1ag/D8, Virtual Bridged Local Area Networks – Connectivity Fault Management*

[11] Nortel Networks, *Presentations: Ethernet Directions PBT Overview, Ethernet OAM IEEE 802.1ag and ITU-T Y.1731, Frame Expansion (802.1as), IEEE-SA Standards Process*, http://www2.nortel.com/go/solution_assoc.jsp?segId=0&parId=0&catId=0&rend_id=17923&contOid=100188013&prod_id=55120&locale=en-US

[12] IEC (Paul Indoo, Nortel), March 2007, *OAM Ethernet*, http://www.iec.org/newsletter/march07_1/broadband_2.html

[13] Nortel Networks, 2006, *Ethernet now offers the most comprehensive OAM for packet-based solutions*, http://www.nortel.com/solutions/collateral/nn119660.pdf

[14] ITU-T, May 2006, *Y.1731, OAM functions and mechanisms for Ethernet based networks*

[15] IEEE Computer Society, December 2005, *Virtual Bridged Local Area Networks*, http://standards.ieee.org/getieee802/download/802.1Q-2005.pdf

[16] IEEE Computer Society, December 2005, *Provider Bridges*, http://standards.ieee.org/getieee802/download/802.1ad-2005.pdf

[17] Nortel Networks, 2006, *Adding Scale, QoS and Operational Simplicity to Ethernet*, http://www.nortel.com/solutions/collateral/nn115500.pdf

[18] Nortel Networks, 2007, *Provider Backbone Bridges Bring Massive Service Scalability to Ethernet*, http://www.nortel.com/solutions/collateral/nn120620.pdf

[19] Worldwide Packets, 2007, *Provider Backbone Transport of Carrier Ethernet Services*, http://www.wwp.com/technology/white-papers/WWP-Provider-Backbone-Transport.pdf

[20] D. Fedyk and L. Andersson, March 2007, *GMPLS Control of Ethernet Forwarding*, http://www3.ietf.org/proceedings/07mar/slides/ccamp-12/ccamp-12.ppt

[21] IEEE 802.1 Working Group, March 2007, *IEEE P802.1ah/D3.4, Virtual Bridged Local Area Networks – Provider Backbone Bridges*

[22] IEEE 802.1 Working Group, May 2007, *IEEE P802.1Qay/D0.0, Virtual Bridged Local Area Networks – Provider Backbone Bridge Traffic Engineering*

[23] ITU-T, *Summary G.8031*, http://www.itu.int/itudoc/itu-t/aap/sg15aap/recaap/g.8031/g8031s_ww9.doc

[24] D. Kent Stevens, 2007, *Carrier Ethernet - A Wave is Building - Provider Backbone Bridges with TrafficEngineering (PBB-TE)*, http://cenic07.cenic.org/program/slides/cenic-2007-kentstevens-nortel.pdf

[25] TPack A/S, 2006, *PBT, Carrier Grade Ethernet Transport*, http://downloads.lightreading.com/wplib/tpack/TPACK_PBT_WP_v1_web.pdf

[26] Nortel Networks, 2006, *PBT, Provider Backbone Transport Achieving true carrier-grade Ethernet*, http://www.nortel.com/solutions/collateral/nn114980.pdf

[27] Nortel Networks, *Productpage Nortel Ethernet Routing Switch 8600*, http://www2.nortel.com/go/product_content.jsp?parId=0&segId=0&catId=-9205&prod_id=44781&locale=en-US

[28] Soapstone Networks, 2007, *Website: Soapstone Networks*, http://www.soapstonenetworks.com/solutions.html

[29] Thomas Nolle, Network World, February 2007, *Will Soapstone end routing's magic kingdom?*, http://www.networkworld.com/columnists/2007/022707nolle.html?page=2

[30] Optical Keyhole, February 2007, *Soapstone Networks - Avici's new open control plane solution business*, http://www.opticalkeyhole.com/interviews/soapstone.asp#s16

[31] TPack A/S, 2007, *Reliable Mobile Backhaul Packet Transport using PBB-TE and T-MPLS*, http://www.tpack.com/publicattachment/white%20papers/COMB_WP_v1_web.pdf

[32] Daniel Joseph Barry, May 2007, *T-MPLS and PBT/PBB-TE offer connection oriented packet transport*, http://lw.pennnet.com/display_article/293047/13/ARTCL/none/none/T-MPLS-and-PBT/PBB-TE-offer-connection-oriented-packet-transport/

[33] Meriton, *The Great Tunnel Debate: PBT vs T-MPLS*, http://www.meriton.com/products/cet-debate.php

[34] Optical Keyhole, April 2007, *Nortel's position on PBB-TE and other issues*, http://www.opticalkeyhole.com/interviews/nortel-PBB-TE.asp

[35] Maarten Vissers, March 2007, *CFM in PBB-TE*, http://www.ieee802.org/1/files/public/docs2007/new-vissers-cfm-in-pbb-te-0307.pdf

List of Figures

List of Tables

Appendix A

Evaluation

This chapter contains an evaluation of this one-month research project. During this research project the subject PBT has been researched. PBT is a new development that gets a lot of attention, therefore a lot of information can be found on the internet. Often the information was superficial, this is due to the fact that IEEE 802.1Qay is currently in draft standard version 0.0.

Other documents contained management summaries only, and did not contain very in-depth information. However, some companies like British Telecom and Shanghai Telecom are experimenting with Provider Backbone Transport. It is expected that more companies and institutions, will test and implement PBT, therefore more valuable information about this subject will be produced.

In the project plan of this research we stated that a Proof of Concept was going to be created. During this research a supervisor of this project tried to obtain a module for the Nortel ERS8600 that was going to be used. This card was obtained in the last week of this research project. Some problems occurred during the installation of the module, therefore it was not possible anymore to test the PBT enabled device.

Still, PBT and its surrounding protocols are part of a complicated environment, which is still part of active development. Therefore it is more difficult to get an overview of the complete protocols and mode of operations of all these protocols, especially because of the time limit of this project. However, we think we managed to get a clear view of PBT, and have found solutions for implementing PBT in combination with SURFnet6.

Appendix B

Glossary

B-DA	Backbone Destination Address
B-SA	Backbone Source Address
B-VID	Backbone VLAN ID
CCM	Continuity Check Message
CFM	Connectivity Fault Management
CPL	Common Photonic Layer
CST	Common Spanning Tree
C-VID	Customer VLAN ID
DA	Destination Address
DRAC	Dynamic Resource Allocation Controller
ERS	Ethernet Routing Switch
GFP	Generic Framing Procedure
IP	Internet Protocol
I-SID	Service Instance Identifier
I-TAG	Service Instance TAG
ITU	International Telecommunications Union
ITU-T	ITU Telecommunication Standardization Sector
LAN	Local Area Network
LBM	LoopBack Message
LBR	LoopBack Reply
LLC	Logical Link Control
LLDP	Link Layer Discovery Protocol
LLDPDU	Link Layer Discovery Protocol Data Unit
LTM	LinkTrace Message
LTR	LinkTrace Reply
MAC	Media Access Control
MAN	Metropolitan Area Network
MD	Maintenance Domain
MEF	Metro Ethernet Forum
MIB	Management Information Base
MPLS	MultiProtocol Label Switching
MSAP	MAC Service Access Point
MSTP	Multiple Spanning Tree Protocol
NOC	Network Operations Center
OAM	Operations, Administration and Maintenance
OME	Optical Multiservice Edge
OPN	Optical Private Network
OTN	Optical Transport Network

PB	Provider Bridges
PBB	Provider Backbone Bridges
PBBN	Provider Backbone Bridges Network
PBB-TE	Provider Backbone Bridges with Traffic Engineering
PBCB	Provider Backbone Core Bridge
PBEB	Provider Backbone Edge Bridge
PBT	Provider Backbone Transport
PDU	Protocol Data Unit
PNC	Provider Network Controller
PPP	Point-to-Point Protocol
QoS	Quality of Service
SA	Source Address
SDH	Synchronous Digital Hierarchy
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
SONET	Synchronous Optical NETworking
STP	Spanning Tree Protocol
S-VID	Service VLAN ID
TDM	Time Division Multiplexing
TLV	Time To Live
TLV	Type, Length and Value
T-MPLS	Transport MultiProtocol Label Switching
TPID	Tag Protocol ID
VID	VLAN ID
VLAN	Virtual Local Area Network
WAN	Wide Area Network