VPLS Technology and Its Applications

Xu Minghai, Mi Zhengkun
(College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, P. R. China)

Abstract:
Among all the metro Ethernet technologies, Virtual Private LAN Service (VPLS), a Layer 2 Virtual Private Network (L2VPN) technology, has drawn wide attention due to its simplicity, reliability, and ease of use. VPLS can implement multipoint-to-multipoint interconnection of Local Area Networks (LAN) in the wide area. Its core elements include the Label Distribution Protocol (LDP) based or Border Gateway Protocol (BGP) based Pseudowire (PW) setup and maintenance on the control plane, Media Access Control (MAC) address learning on the data plane, and PW encapsulation on the transport plane. With hierarchical architecture, VPLS enables cross-domain Virtual Local Area Network (VLAN) services. Benefitting from its unique technical advantages, VPLS supports a wide variety of applications including L2VPN for VIP customers, municipal communications infrastructure and personal distributed services.

One of the important evolution trends for Ethernet is to use the Multi-Protocol Label Switching (MPLS) technology for MAN to carry Ethernet data frames, so as to provide virtual Ethernet services that enables interconnection of geographically dispersed LANs[1-2].

IP and Ethernet have run neck and neck in the investment race of global telecom industry in recent years, and Ethernet over MPLS (EoMPLS) outshines the rest in growth[3]. Furthermore, standards organizations such as IEEE, IETF, ITU-T and Metro Ethernet Forum (MEF) have conducted an in-depth research on Ethernet and its derivative technologies.

Provided by the Packet Switching Network (PSN), Virtual Private LAN Service (VPLS) aims to offer dedicated LAN interconnection service through Pseudowire (PW) connection in pre-established tunnels and falls into the category of Layer 2 Virtual Private Network (L2VPN). Theoretically, VPLS can use any type of tunnels, such as MPLS Label Switching Path (LSP), Generic Routing Encapsulation (GRE), Layer 2 Tunneling Protocol (L2TP) and Internet Protocol Security (IPsec) extensions, among which the MPLS LSP is most widely used. As a significant technology of EoMPLS, VPLS integrates advantages of IP/MPLS, Virtual Private Network (VPN) and Ethernet switching, and supports multipoint-to-multipoint interconnection of LANs across wide areas. For carriers, L2VPN is simple and transparent, and can lower network complexity and enhance network interoperability.

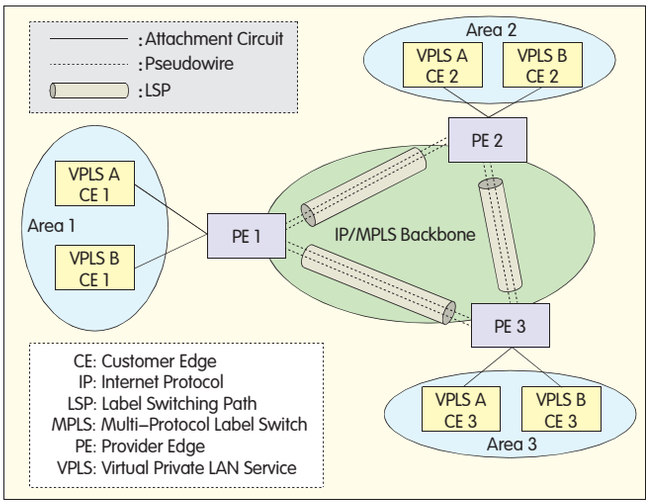
This article focuses on VPLS technology that adopts MPLS LSP as the transport tunnel and Label Distribution Protocol (LDP) as the PW setup signaling, and its application.

1 VPLS Network Architecture and Core Technologies

1.1 VPLS Network Architecture

Figure 1 shows the basic network architecture of VPLS, which assumes that each of two corporate customers A and B has three branches that are respectively located in Areas 1, 2 and 3. To interconnect LANs located in these three areas, the carrier configures three devices called Provider Edge (PE) devices respectively in Areas 1, 2 and 3. Correspondingly, customers have devices directly connected with the carrier networks in each LAN. These devices are called Customer Edge (CE) devices, which are connected with corresponding PE 1, PE 2 and PE 3 through the Attachment Circuits (AC). The types of AC are independent of VPLS and can be physical/logical Ethernet port, frame relay link, Asynchronous Transfer Mode (ATM) Permanent Virtual Connection (PVC) or even Ethernet PW[4-5]. All PEs interconnect through LSPs. As shown in Figure 1, two PWs are set up in each LSP, serving customers A and B respectively. PW refers to a bidirectional emulated point-to-point connection between two nodes, and consists of two unidirectional LSPs. The carrier transports data streams among LANs in different areas over public PSN through PE and interconnected PWs, and thus interconnects several geographically dispersed LANs as a single emulated LAN, known as a VPLS

This work was supported by the Startup Program of Innovative Young Talents of Jiangsu Provincial Natural Science Foundation under Grant No. BK2007604.



▲Figure 1. VPLS network architecture.

instance. The LAN in each area can be deemed as a network segment of this emulated LAN. The IP/MPLS backbone network may consist of one or several interconnected domains and spans a wide area, so the carrier can provide LAN interconnection service that spans MAN or WAN by the use of VPLS. From customers' perspective, VPLS helps to simplify networking and eliminates the need to change existing network architectures.

As shown in Figure 1, similar to the function of multiplexer in a transmission network, the tunnel between PEs carries the aggregated traffic flow of several VPLS instances. The specific protocol adopted for tunnel setup is subject to the tunneling technology applied, for example, in an MPLS network, the LDP can be adopted for LSP setup. Equal to a demultiplexer, the PW in a LSP carries traffic flow of a single VPLS instance. Each PW is assigned with a Pseudowire Denotation (PWD), which can be seen as a PW identifier. A VPLS instance can be implemented by setting up PW connections between PEs, which need to define related PW control signaling by extending the use of LSP setup signaling protocol. Upon the establishment of VPLS instance, PE functions as an emulated Ethernet bridge to forward Ethernet frames to corresponding PWs and then to the destination customer LAN, while these frames are transmitted to the PE from a LAN located in certain other areas belonging to the same customer through AC so as to implement LAN interconnection across different areas.

VPLS primarily involves the following three technologies:

(1) Control plane technology: PW control signaling is used to set up and maintain PW connections[6–8], which is responsible for PW setup and teardown, notification of PW status changes, and PW protection.

(2) Data plane technology: It primarily involves the bridging and forwarding function of PE, especially the MAC address learning.

(3) Transport plane technology: It is PW encapsulation type of VPLS packets.

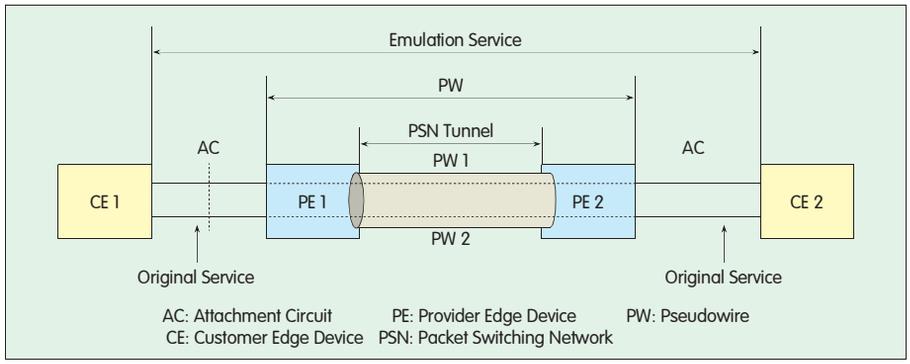
1.2 LDP–Based PW Setup

Figure 2 shows the network reference model of Pseudowire Emulation Edge–to–Edge (PW3E)[9]. For ease of illustration, suppose PE 1 is an ingress router and PE 2 is an egress router. An MPLS LSP is established between PE 1 and PE 2, and multiple PWs can be carried in the LSP; each PW corresponds to one AC between CE and PE; PW emulation transmits the original service carried on related ACs; the original service can be any of the types such as frame relay, ATM, Ethernet and IP. The PW connection is set up by adopting appropriate signaling protocol to realize negotiation between PE1 and PE2 and establish the binding of PWs and LSPs as well as the binding of PWs and ACs on related ports.

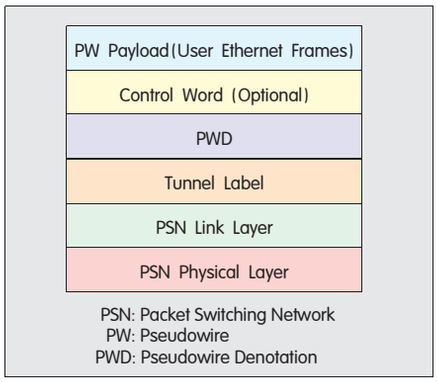
If the PSN type is MPLS, LDP can be further used as the PW setup and maintenance signaling protocol. The core concept of LDP is to establish binding between MPLS LSP label and Forwarding Equivalence Class (FEC). Therefore LDP can be used directly to establish binding between PWs and LSPs as long as PWs are treated as one of FECs.

For the reasons above, the LDP extension defines two types of PW FECs: PWid and generalized PWid. The parameter PWtype in PWid defines the PW type, and generally corresponds to original services that are emulated in the PW3E network, such as frame relay, ATM, High–Level Data Link Control (HDLC), Point–to–Point Protocol (PPP) and Synchronous Digital Hierarchy (SDH). The parameter PWid specifies a PW of the type, that is, to locate ACs. For the same PW, identical PWtype and PWid must be configured for both PE 1 and PE 2. Apart from the parameter PWtype, the generalized PWid uses Attachment Group Indicator (AGI), Source Attachment Individual Identifier (SAII) and Target Attachment Individual Identifier (TAII) to locate PW related ACs. Where the AGI indicates AC group, such as VPN and VLAN identification; the AII is the Attachment Individual Identifier, used to locate certain AC in the AC group; the SAII and TAII respectively indicate source (local end) and target (remote end) ACs. For VPLS applications, the AGI can be set to the VPLS identifier. PWs are meshed and are not used for forwarding, and AC can be located only by the VPLS identifier, so SAII and TAII are unnecessary and generally set to zero.

The LDP–based PW setup adopts the unsolicited downstream label distribution mode and liberal label retention mode,



▲Figure 2. PW3E network reference model.



▲Figure 3. PW encapsulation format.

that is, the egress router PE 2 voluntarily initiates PW setup procedure by sending a Label Mapping Message (LMM) to the ingress router PE 1. The LMM contains two information elements: One element is PW FEC, which can be PWid or generalized PWid; the other element is LSP label, which adopts the 20-bit generalized label defined in MPLS to realize the binding of MPLS LSPs and PWs. Upon receiving the LMM, PE 1 checks whether the PW LSP in reverse direction (PE 2→PE 1) is set up; if not, it sends an LMM to PE 2 to establish bidirectional PW connection in a similar way. If PE 1 fails to identify the PW FEC in the LMM, it returns a Label Release Message (LRM) to reject the request.

After PW setup, PEs can use LDP to exchange PW status, perform PW maintenance and inform other PEs to delete learned MAC addresses so as to speed up forwarding table convergence.

1.3 PW Encapsulation

In the PW encapsulation format shown in Figure 3, PW payload refers to user Ethernet frames transmitted on the PW. The optional control word contains a 16-bit sequence number used by peer PE to check whether transmitted Ethernet frames are disordered, repeated or lost. The need for the control word in the encapsulation depends on the specific application requirement and is indicated by the PW FEC in the LMM. The PWD is used for PW identification. The LSP label identifies MPLS LSP encapsulation. The link layer and physical layer of MPLS-based PSN are at the bottom.

After user Ethernet frames reach PE through AC, PE first removes the header scrambling sequence and frame check sequence, and then performs appropriate processing for possible labels of Ethernet frames based on PW types. There are two PW types in VPLS application: Ethernet PWs and Ethernet VLAN PWs. The specific PW type is indicated in PW FEC. An Ethernet PW handles labels in raw mode, and its processing rules are as follows:

(1) If user Ethernet frames contain service-delimiting tag (which is used to identify services of different customers or services of the same customer with different QoSs) set by the carrier, for example, VPLS instance identifier, ingress PE will remove the tag and egress PE may or may not re-insert service-delimiting tag depending on different requirements.

(2) If user Ethernet frames contain customer-set tag, for example, VLAN tag, PE will reserve the tag.

Therefore, raw user Ethernet frames are transmitted on Ethernet PWs at all times. For an Ethernet VLAN PW operating in tagged mode, the processing rules are as follows:

(1) If user Ethernet frames contain service-delimiting tag set by the carrier, for example, VLAN tag to indicate a VPLS instance, PE will reserve the tag, but if an identifier has been specified during PW establishment, this identifier must be used to replace the existing VLAN tag.

(2) If user Ethernet frames contain no service-delimiting tag, PE must add the tag specified during PW setup, and if no tag is specified, PE needs to add an empty tag.

Therefore, user Ethernet frames transmitted on Ethernet VLAN PWs always contain service-delimiting tags.

1.4 MAC Address Learning

PE functions as a Virtual Switching Instance (VSI) on data plane and needs to implement filtering, learning and forwarding functions as a bridge.

- Filtering: To check and analyze transmitted Ethernet frames headers.
- Learning: To establish forwarding table entry by analyzing MAC address in frame header, that is, to establish a binding relation between MAC addresses and PWs or ACs.
- Forwarding: To search in the forwarding table with the destination MAC address of Ethernet frames, and forward unicast Ethernet frames to PWs or ACs bound to exsiting MAC addresses; to broadcast unicast frames with destination MAC address unknown or broadcast frames to all PWs.

The method of MAC address learning is described as follows:

(1) If there is no forwarding entry corresponding to the source MAC address of an Ethernet frame received by PE from AC, PE will bind the MAC address with the AC so that it can forward any Ethernet frame to AC when its destination MAC address matches the bound MAC address.

(2) If there is no forwarding entry corresponding to the source MAC address of an Ethernet frame received by PE from PW, PE will bind the MAC address with the PW so that it can forward any Ethernet frame to PW when its destination MAC address matches the bound MAC address.

(3) Each MAC address entry is set with an aging timer. If no Ethernet frame is transmitted using its bound MAC address until the timer expires, this obsolete MAC address entry will be deleted, which is called address aging mechanism.

(4) In the case of change in PW or AC status, PE needs to re-operate the bridge learning mechanism.

One of the basic requirements for bridge forwarding is to avoid loops. The Spanning Tree Protocol (STP) is used in Ethernet to avoid loops, but STP has a drawback of slow convergence. To accelerate the convergence, VPLS adopts the topology structure in which all PEs are fully meshed to ensure network connectivity. The frame forwarding complies with the "Split-Horizon" rule, that is, no PE is allowed to forward frames received from PWs to other PEs through PWs to prevent routing loops.

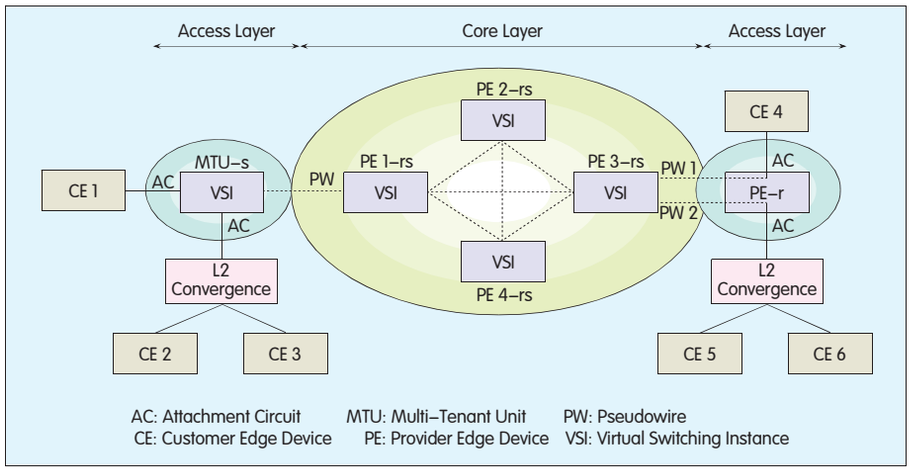
2 Hierarchical VPLS

In the above-mentioned topology, all PEs are fully meshed, for a VPLS instance, this equation exists: The number of PWs = $n \times (n-1)/2$; "n" refers to the number of PEs. When the VPLS network is large in scale, a substantial amount of network resources, signaling overhead and data frame duplicating/processing overhead is

required, thus restricting network scalability. Therefore, the Hierarchical VPLS (HVPLS) technology is introduced to divide a VPLS network into two layers: access layer and core layer. Specifically speaking, the devices at access layer are close to customers; they converge traffic flow of several customer areas (CEs) and send it to PE at the core layer through connected PWs, namely, spoke PWs, thus reducing the number of meshed PWs at the core layer.

The HVPLS network, as shown in Figure 4, contains two types of access layer devices: Multi-Tenant Unit-switch (MTU-s) and Provider Edge-router (PE-r). MTU-s, with built-in VSI, supports Layer 2 switching as well as bridge learning and frame duplication functions on all ports. CE 1, CE 2 and CE 3 are three sites of a certain VPLS instance. The VSI of MTU-s implements data interaction among them without sending data to the core layer. The data stream from CE to peer end is sent to PE1-rs (rs stands for routing and switching) in the core layer after being converged by VSI. MTU-s is capable of switching, so only one PW is required for each VPLS instance. As shown in Figure 4, although each CE accesses MTU-s through two ACs, yet only one spoke PW is required between MTU-s and PE1-rs. MTU-s may also connect with several PEs at core layer to improve reliability. PE-r supports only routing function and is not capable of bridging. So it is required to setup a PW to directly connect the PE-r to the core PE for each access AC and data frames from AC will be routed to related PWs, so data stream between CE 4 and CE 5 or CE 6 must be forwarded through PE3-rs. PE-rs devices at the core layer are fully meshed and are capable of both routing and switching. The introduction of access layer devices reduces the number of fully meshed PE-rs devices and thus enhances VPLS scalability.

Further, the above access layer structure in which access devices are connected to the core layer through spoke PWs can also be extended into Ethernet access network. At present, a lot of Ethernet access networks support VPLS, and they identify VPLS instances by adding VLAN tag set by the carriers. In this way, HVPLS is extended into a



▲Figure 4. HVPLS network model.

two-layer network consist of access-layer VPLS and core-layer VPLS. Usually both MTU-s and PE-r are called User PE (U-PE), and PE-rs at the core layer is called Network PE (N-PE). U-PE and N-PE are no longer directly connected through spoke PWs; instead, they are connected by access VPLS network based on IEEE 802.1ad (Q in Q) or MPLS technology. In this way, up to several thousands of VPLS customer sites can be converged to access the core network through access networks. Thus the core network of the same scale can provide VPLS services for more customers, which will effectively enhance VPLS scalability.

3 Application of VPLS

The past few decades have witnessed swift development and wide application of Ethernet technologies as well as constant drop of Ethernet deployment cost. As an extension from Ethernet to MAN/WAN, VPLS combines the advantages of network performance and network scale and offers new options for network and service operation. Carriers can flexibly deploy VPLS based on customer types and service attributes. Owing to its unique technical advantages, VPLS can be widely applied in such scenarios as L2VPN for VIP customers, VPLS-based metropolitan communication infrastructure and individual distributed services.

3.1 L2VPN for VIP Customers

This is the most used application of VPLS. Compared with L3VPN, VPLS features the following technical advantages:

(1) The user network is relatively independent of the carrier network. The latter offers data isolation and transparent transmission for user data, so as to ensure user data security and avoid impacts brought by the complexity of the carrier network.

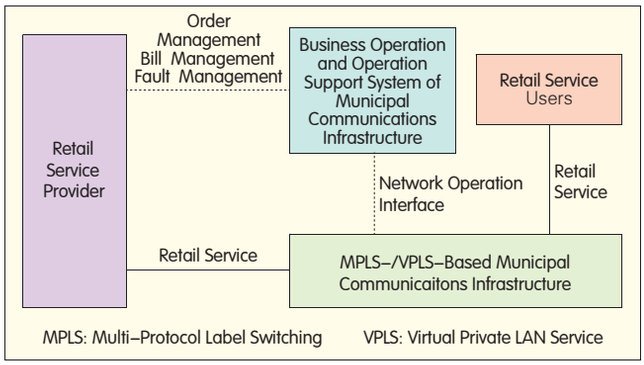
(2) The implementation of Ethernet multipoint services in the MPLS network can complement Ethernet technologies with the advantages of MPLS technology.

(3) VPLS features technical simplicity/reliability, deployment flexibility, and excellent scalability.

Corporate customers may have branches spread across the whole province, nation and even the globe, and therefore, it is a wise choice to perform L2VPN connection of all branches by use of VPLS/HVPLS. When a VPLS network is large (with a huge amount of geographically dispersed nodes), HVPLS that combines Border Gateway Protocol (BGP)-based VPLS and LDP-based VPLS can be adopted; BGP is used at the core layer while LDP is used at the access layer.

3.2 VPLS-Based Municipal Communications Infrastructure

An operable and manageable municipal communication network should become a component of municipal infrastructure like water, electricity and gas networks. Figure 5 shows the reference model of VPLS-based municipal communications infrastructure[10]. The business operation



◀Figure 5. Operation reference model of VPLS–based municipal communications infrastructure.

and operation support system is responsible for the construction, operation and maintenance of MPLS–/VPLS–based municipal communications infrastructure. The retail users sign service contracts with retail service providers. The retail service providers pay traffic fees for the business operation and operation support system.

MPLS/VPLS–based municipal communications infrastructure may involve multiple layers, for example, subscriber layer, access layer, distribution layer, local core layer, regional core layer, provider access layer and provider layer. The subscriber layer is responsible for user access. The access and distribution layers converge user traffic. The local and regional core layers transmit converged traffic on high–speed backbone links. The service provider accesses to the core layer through the provider layer and provider access layer. MPLS/VPLS connects switches at the core layer (including local and regional core layers) with provider–oriented switches. The network creates an VPLS instance for each service, and specifies QoS level by using the EXP field in the MPLS label.

3.3 Personal Distributed Services

Customer–centered communications mode will become a promising application along with social development and progress, it interconnects Personal Area Network (PAN), Home Network (HN) and Office Network (ON) through MAN/WAN, so that users can control devices and sessions of all these subnets[11–14].

In this mode, three ACs can be used to connect PAN, HN and ON gateways to MAN PE respectively; PE creates VPLS instances for users, and connects the forwarders of VPLS instances through PWs. From users' perspective, the VPLS network functions as a "distributed" Ethernet switch. By the use of VPLS, users can conveniently control subnet devices and create sessions, and realize secure and reliable communications anytime anywhere.

4 Conclusions

By making full use of already deployed MPLS networks to provide LAN interconnection services for VIP customers and functioning as MAN infrastructure, VPLS has been considered to be one of the significant carrier–class Ethernet technologies with broad market prospects. VPLS can be further explored in the following aspects:

- Enhance VPLS reliability by adopting redundancy and multi–homing mechanism;
- Improve VPLS network resource utilization through effective resource control with guaranteed QoS;
- Construct a service bearer model for overlay networks to enhance VPLS capabilities for carrying Value–Added Services (VASs) and converged services.

Finally, it must be noted that VPLS and other metro Ethernet technologies are complementary, and therefore an optimal result can be achieved only by making appropriate selection and combination of these technologies based on actual conditions.

References

[1] MALIS A G. Converged services over MPLS [J]. IEEE Communications Magazine, 2006, 44(9): 150–156.
[2] ALLAN D, BRAGG N, McGUIRE A, et al. Ethernet as carrier transport infrastructure [J]. IEEE Communications Magazine, 2006, 44(2): 95–101.
[3] SEERY M. Packet transport trends: IP/MPLS Success challenged as deployment footprint expands [J]. IEEE Communications Magazine, 2008, 46(7):103–107.
[4] ANDERSSON L, ROSEN E. Framework for layer 2 virtual private networks (L2VPNs) [J]. IETF RFC 4664. 2006.
[5] ANDERSSON L, MADSEN T. Provider provisioned virtual private network (VPN) terminology [J]. IETF RFC 4026.2005.
[6] BRYANT S, PATE P. Pseudo Wire Emulation Edge–to–Edge (PWE3) Architecture [S]. IETF RFC 3985, 2005.
[7] MARTINI L, ROSEN E, EL–AAWAR N, et al. Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP) [S]. IETF RFC 4447. 2006.
[8] LASSERRE M, KOMPELLA V. Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling [S]. IETF RFC 4762. 2007.
[9] MARTINI L, ROSEN E, EL–AAWAR N, et al. Encapsulation Methods for Transport of Ethernet over MPLS Networks [S]. IETF RFC 4448. 2006.
[10] MOERMAN K, FISHBURN J, LASSERRE M, et al. Utah's UTOPIA: an Ethernet–based MPLS/VPLS triple play deployment [J]. IEEE Communications Magazine, 2005, 43(11): 142–150.
[11] ATKINSON R C, IRVINE J, DUNLOP J, et al. The personal distributed environment [J]. IEEE Wireless Communications, 2007, 14(2): 62–69.
[12] AUSUSTYN W, SERBEST Y. Service Requirements for Layer 2 Provider–Provisioned Virtual Private Networks (L2VPNs) [S]. IETF RFC 4665. 2006.
[13] ITU–T Draft Recommendation Y.enet. Ethernet QoS control for next generation networks [R]. Geneva, 2008.
[14] BOCCI M, COWBURN I, GUILLET J. Network high availability for Ethernet services using IP/MPLS networks [J]. IEEE Communications Magazine, 2008, 46(3): 90–96.

Biographies

Xu Minghai

Xu Minghai, PhD, is a lecturer at the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications (NUPT). His research interests cover resource control and QoS, mobility management, and network recognition technology. He has led 2 provincial level scientific research projects, and participated in 1 national "973" project and 1 national natural science foundation project. He has submitted 6 proposals for international communication standards and 2 patent applications, and published 8 research papers indexed by SCI/EI/ISTP.

Mi Zhengkun

Mi Zhengkun is a professor at the College of Telecommunications and Information Engineering, NUPT. He is also a fellow of China Institute of Communications. He focuses his research on Next Generation Networks (NGNs) and heterogeneous network convergence technologies. Owing to his research achievements, he won the second prize of Scientific and Technological Progress Award granted by Jiangsu province, and the second and third prizes of Scientific and Technological Progress Award by Ministry of Industry and Information Technology of the People's Republic of China. He has published more than 30 papers indexed by SCI and EI, 8 monographs and state–level textbooks, and applied for more than 10 national patents.