2004－AVM－47（4）
2004／12／9

MPsLS: A Forwarding Scheme of Guaranteeing QoS in Integrated Services Networks

Jun YANG　　Yasushi HIBINO
School of Information Science, Japan Advanced Institute of Science and Technology
Address: 1-1, Asahidai, Tatsunokuchi, Ishikawa 923-1292 Japan
E-mail: j-yang, hibino @jaist.ac.jp

Abstract

In this article, a novel data forwarding scheme for a core area of the integrated data service network is proposed. The scheme is named MPsLS, in which synchronous frames are used to transfer data. Appointed channels in the frame transmit time-sensitive application data in synchronous mode, at the same time, filler channels in the frame transmit non time-sensitive application data in pseudo asynchronous mode. Therefore, MPsLS can guarantee QoS completely for time-sensitive applications, and maintain the high utilization of network resources.

Key words: QoS, synchronization.

1. Introduction

Nowadays, the information networks are separately existing for the real-time application (such as telephone) and for the non real-time application (such as Internet) due to their specific functions. This situation obviously results in huge wastage of network equipment resource. The reason of separation is that the ideal switching technologies for the real-time network is different for the non real-time network.

The switching technology for the real-time information network is the circuit switching, which can inherently guarantee requited QoS to transport data because the traffic from an application is completely separated from other ones throughout the network, therefore there will occur no congestion or delay variation. The typical real-time network is the current SDH/SONET whose main feature is Time Division Multiplexing (TDM)[1-4] technology to provide multiple channels by dividing bandwidth, However, traditional TDM technology also has a number of drawbacks, for example, static nature of connections, lack of dynamics for resource distribution, and expensive manual management.

There is a method to improve channel utilization, called Dynamic synchronous Transfer Mode (DTM)[5] which extends TDM by using dynamic multiplexing scheme and dynamic channels allocation at the lowest level based on current SDH/SONET technology. However, in this scheme, when the bit rate of a application traffic is largely varying, the resource utilization becomes low.

On the other hand, non real-time information networks apply the packet switching technology, which has brought huge developments because of its good flexibility. However, the packet switching technology can not guarantee QoS requirement to deliver data for the specific application. It is due to the fact that IP connectivity only provides best-effort services.

Therefore, development of integrated technology for seeking both QoS guarantee and high resource utilization is more and more strongly urged.

Recently, as advances of optical network technology, Generalized MPLS(GMPLS)[6-8], an extension of multi-protocol label switching concept, has been proposed. It will be the future integrated standard of network transfer technology, which has Packet-Switch Capability (PSC), Time-Division Multiplex (TDM) Capability, Lambda (wavelength or waveband) Switch Capability (LSC), and Fiber-Switch Capability (FSC). GMPLS, in addition, has the control plane separating from data plane.TDM layer operates between packet services and the wavelengths as figure 1.

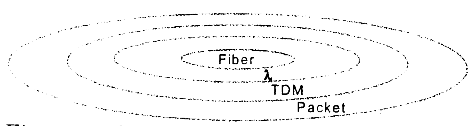

Figure 1. The Layers of Network Structure

Although it is a promising scheme, the suite of GMPLS protocols is under standardization process.

In this paper, a novel forwarding scheme, MPsLS, is proposed as a basis of GMPLS implementation.

2. Basic Principles of MPsLS

The Multi-Protocol synchronization Label Switching (MPsLS) is an extension of MPLS concept that provides interface between the layer-3, the network layer, and the layer-2, the data link layer. MPsLP introduces the frame base synchronous transfer mode with affinity for TDM into layer-2. MPsLP absorbs differences between layer-3 packet transfer and layer-2 synchronous frame transfer based on multi-protocol label switching. MPsLS separates the application traffics into two types, time-sensitive and non time-sensitive. The former are transmitted by appointed channels that consist of fixed position slots of a frame, and the latter are transmitted with filler channels that are formed with plesio-contiguous floating position slots of a frame.

2.1 Data Formats of Synchronization Frame in MPsLS

In MPsLS network, data are transferred in cycle mode with a constant period. During each cycle, the data are transferred continuously. These data in a cycle is called a frame. The size of the frames is depends on the link rate of a fiber or a lambda link. Each frame is divided into a number of slots of same size. The slots are classified into two types, control slots and data slots. The control slots are placed on the header part of the frames, which carry the network control information. The data slots carry the application data. And each data slot usually carries only a segment of application data stream or layer-3 packet. The structure of transfer mode is shown in figure 2.

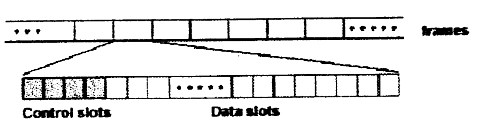

Figure 2: The Structure of the Frame

The data slots are divided into two classes, appointed channels and filler channels, according to differences of carried data traffics, the appointed channels are slots carrying time sensitive traffic, and the filler channels are slots carrying non time sensitive traffic. Each appointed channel can contain several slots according to bandwidth requirement. Their slot positions in a frame are fixed during the channel is being held. Each slot in a filler channel carries only a segment of a layer-3 packet. The number of slots in a filler channel depends on the size of a packet. The filler channel carrying non time-sensitive application data. Their slot positions are floating in different frames.

2.2 Appointed Channels Transfer

In MPsLS network, time-sensitive traffics are synchronously transmitted by appointed channels on same slot position. The structure of an appointed channel slot is that the first two bits are slot type bits, the next 6 bits are shared channels index for channel sharing, the rest bits are payload.

2.3 Filler Channels Transfer

Non time-sensitive traffics are forwarded by filler channels. The slot positions of a filler channel are not fixed during a channel is being preserved.

When there is no time-sensitive traffics to be transmitted, slots can temporarily be used to carry non time-sensitive traffics as a filler channel. The structure of a slot for filler channels is that the first two bits are slot type bits, the rest bits are payload. The first slot of a filler channel includes multi-protocol label switching header and the rest slots are placed on plesio-contiguous position in frames.

2.4 Classes of Providing Services

MPsLS provides three classes of services, Completely Guaranteed Service(CGS), Non-Completely Guaranteed Service(NGS) and Best-Effort Service(BES). For Constant Bit Rate(CBR) applications and real time Variable Bit Rate(rVBR) applications, CGS is applied. Bandwidth for this class requires the peak bit rate.

Since the appointed channel is over-preserved for general VBR, for loss allowable real-time application, NGS is provided. For NGS, the appointed channel is shared by several share channels giving the mean bandwidth required for each shared channels.

NGS will accomplish both good channel utilization and real-time application requirement, constant delay, despite of data loss.

For non real-time application, BES is applied. BES does not guarantee time constraints and data integrity.

3. Control in MPsLS

The network service needs control mechanisms to operate the system running harmoniously. Control messages in MPsLS are organized with three parts. The first is the signaling which is sent from ingress edge nodes in an MPsLS cloud, such as path establishing and releasing messages, these messages are transferred by control slots in a form of appointed channels. The second is the control database these reside in each MPsLS nodes. The third is control bits these occupy the header of each slot.

3.1 Signaling

The signaling information is transmitted by control slots in a form of appointed channels, which include routing messages and channel management messages.

3.1.1 Path Establishment Signaling

Since MPsLS provides three different transfer modes for variety of applications requires different QoS, methods for pat setup are multiple.

When a time-sensitive traffic arrives to an edge node, the edge node send a request message, which includes commands to find if there is or not a free slot position of a frame within allowable offset along the path as an appointed channel. If free slot positions along the route for an appointed channel are found, the edge node reserves the slot positions along the rout for the appointed channel, otherwise, it refuses the request.

The setup process is that, when the message passes through hops, it gathers the information of the free slot position and alter the slot type field to used sate, when the message reaches a destination node, then the message backs to the first edge node along the path, and establish a new path and a appointed channel. In the above new path finding process, if the message can not reach the destination node, it backs to the first edge node maintaining the slot type field to idle state to be usable for other new requests with the gathered free slots position information, and it refuses the request.

For non time-sensitive traffic, according to the admission control protocol, if the bandwidth condition is satisfied, the path for a filler channel can be set up.

3.2 Databases in Nodes

Each node of MPsLS network has several databases to store the control information.

3.2.1 Appointed Channel Switching Table

The appointed channel table is set up for each input port. It tells the in-coming frame the output port for each appointed channel and the position of the out-going frame. The appointed channels table is only used for appointed channels switching. Operation is carried on table-driven manner in every cycle. In each cycle, the node firstly examines the slot type field. If a slot is carrying time-sensitive traffics, then refers the appointed channel table to find a corresponding output port and the slot position in the out-going frame.

3.2.2 Label Switching Table

Non time-sensitive traffic is transmitted by filler channels. In in order to manage the filler channel data forwarding, a label switching table is provided for a node, It is updated when a virtual path is set up. For each arrived packet on a filler channel, the node examines the header of the filler channel and refers to label switching table, and finds outgoing port, then swaps the label if necessary and forwards the packet of the filler channel to a corresponding wait queue to be transmitted.

3.2.3 Appointed Sub-Channel Table

Edge routers of MPsLS have another table besides the two tables above, the appointed sub-channel table, which keeps sub-channels information of sharing an appointed channel. (See. 4.5)

3.3 Slot Types Field

The first two bits of a slot in a frame indicate types of data contained in the slot. When the value is 11, the slot carries time-sensitive data traffic which is preserved, when the value 10 means loss allowable time-sensitive traffic data. 01 is for a filler channel and 00 means the slot is free.

4. Process in Ingress and Egress Nodes

Edge nodes of MPsLS not only have the functions of constructing frames, but provide more ones than in core nodes, such as sending out and receiving path request massage, deciding the types of the services, and segmenting and reassembling the packets.

4.1 Types of Applications

The applications in integrated digital networks, traffics are classified into three types, 1) guaranteed time-sensitive applications, 2) loss allowable time-sensitive applications, and 3) non time-sensitive applications, the last are insensitive to the delay and jitter, such as file transfers and E-mail service. The former two are sensitive to the delay and jitter, such as live audio and video data transmission. It is possible to lead to noticeable degradation of service quality with the increase of delay and jitter. Thus guaranteeing quality of service (QoS) is needed for this class of applications. The second is time-sensitive applications that allow data loss partly.

4.2 Establishing Paths

When the traffic from a user arrives to an ingress edge node, the node sends a path set up request message on a control slot in the form of appointed channel, the message is transferred hop by hop, to an egress node. If all resources and time requirement can been satisfied, the path is set up, otherwise, the request is rejected.

4.3 Segmentation and Reassembly of Packets

All information in the IP networks, is transferred in the form of packets, and in the MPsLS cloud, information is transmitted on time slots of fixed length.

Since lengths of most packets are longer than a slot, segmentation and reassembly process is inevitable at an ingress and an egress edge node of the MPsLS cloud.

Furthermore, if users want to connect an MPsLS cloud with a special MAN based on such as DTM, in order to reduce the overhead caused by packetization, a bit stream can be directly segmented into slots at an ingress, and slots can also be reassembled into a bit stream at an egress node.

4.4 Constructing Synchronization Frames

At first, the ingress node segments the guaranteed time-sensitive applications packets and loss allowable time-sensitive applications packets into data pieces and fill corresponding slots in preserved positions for appointed channels, then write the type of slots, 11 in the first two bits of a corresponding slot, or 10 expresses for a loss allowable slot. After finishing all time-sensitive applications packets, non time-sensitive applications packets are handed. The first two bits indicate the type of slots, the next 32 bits of the first slot are label switching header, the rests are payload. If a slot is free, it is marked with 00.

4.5 Sharing Appointed channel

Since bandwidth of each slot is 2Mbps when the total bandwidth of a lambda is 10 Gbps, a frame time is 125 mu sec and the number of slots is 1000, if required bandwidth of an application is less than 2Mbps, a part of bandwidth of a slot is wasted.

To overcome the shortcoming, the authors use bundle technology with several sub-channels, namely, several low bit rate real-time application streams share same appointed channel when the both edge routers(ingress and egress) are same. The authors set the least sub-appointed channel is 64Kbps. It means that every 8 bits of a slot can carries a independent application. After a appointed channel is established, if left bandwidth is greater than 64Kbps, the related information is saved in the appointed sub-channel table. When there is another application with same edge routers, its required bandwidth is equal or less than the left bandwidth, the new application send a sharing request message, and the message arrives the destination node, the destination node returns the permission message, and at the same time, all nodes establish the sharing appointed channel tables.

This technology makes several small bandwidth bit stream share a appointed channel. However, when required bandwidth is greater than 2Mbps, it had better to assign a new slot for the appointed channel, which can reduce the overhead caused by segmenting different length of data pieces for one application.

5. Forwarding at Core Nodes

Data forwarding in MPsLS is based on synchronization. The synchronization processes in each node separate in three stages, namely, input, processing and output synchronization.

5.1 Input Synchronization

Input synchronization is accomplished with a delay line. The difference of propagation delay time among core nodes are large because the difference of the distances among the core nodes in the Internet. Therefore there exist noticeable fiber buffer effect. To fulfill synchronization of input frames, a delay line composite of a fiber loop and a shift register are used. The fiber loop can absorb the delay less than one clock. The shift register adjust more than one clock delay. All frames arrived to the node ports are completely synchronized as if they arrive at same time.

5.2 Process Synchronization

The processes of frames include identifying types of slots, scanning the slots in a frame, looking up appointed channel tables and label switching tables, assembling new frames. These implementation can be finished within several synchronization cycle. These are shown in figure 3.

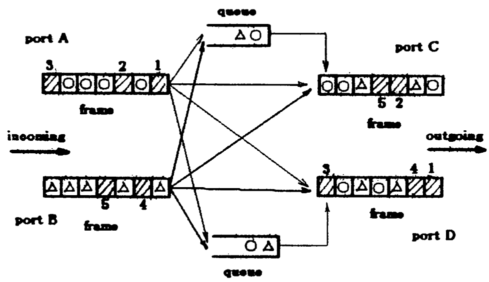

Figure 3. The Process of Frame Forwarding

Figure 3 shows that when incoming frames arrive input buffers. In input buffer of the ingress, the slots are handled one by one, if the slot is appointed channel. They will be sent to the position with allowable offset of corresponding outgoing frame according to the information of the appointed channel table. If the slots are filler channel, the data in slots are sent to the corresponding queue according to the label switching table. At the same time, in outgoing frames, if no slots receive as a appointed channel, slots of a filler channel are read out from the corresponding queue. This process continues to the end of outgoing frame.

5.3 Output Synchronization

When the assembly of the frame is finished, the finished frame on the output buffer will be sent out in next cycle, and the new frame for the next cycle is prepared on the alternate output buffer.

6. Performance Analysis

Usually, three parameters are selected to evaluate performance of transmission scheme. They are time (delay, jitter), loss ratio (data loss, calling loss) and utilization of resources. The performance analysis will be done separately for the every transfer mode with these parameters, because MPsLS provides variety of transfer mode for different application traffics.

6.1 Appointed channel Transfer

In appointed channel transfer mode, since the channels are synchronous, the jitter of delay is 0, the delay is constant which depends on propagating, transmitting, processing time, allowable offset, and cycle period.

For both CBR and VBR applications, the transmission appointed channel are preserved with peak bit rate, therefore, the loss during transmission is 0.

In addition, for same position forwarding mode along the path (allowable offset is zero), we assume that the bottleneck of network occurs in this core area, where the total numbers of appointed channel are same, are equal to m, and used appointed channels are independent with one another, the number of time-sensitive application traffics is equal to k in each segment, then if the hops along the path is n, then the probability of i slots being used, Pi, is

$$P_{i,used} = \begin{cases} 1 & (i < k) \\ P_i & (k \leqslant i \leqslant \min(m, n*k)) \\ 0 & (i > \min(m, n*k)) \end{cases}$$

where,

$$P_i = \frac{{}_iC_m * S_i}{({}_kC_m)^n} \qquad i=(k, \min(m, n*k))$$

$$S_k = ({}_kC_k)^n = 1$$

$$S_i = ({}_kC_i)^n - \sum_{j=k}^{i-1} {}_jC_i * S_j \qquad (i>k>0)$$

and being equal to and more than x free slots, the probability is,

$$P_{x+,free} = 1 - \sum_{i=k}^{m-x} P_{i,used}$$

For a one-slot appointed channel application, the calling loss ratio is shown in table 1.

Table 1. The Relation of Calling Loss Ratio with Ratio of Time-sensitive Parts to all Traffics

calling loss ratio \ ratio; hops n=5		k/m				
		0.5	0.6	0.7	0.8	0.9
m	100	3.113%	34.270%	78.139%	96.841%	99.999%
	1000	0.000%	0.002%	8.473%	72.537%	99.005%

Table 1 shows that when the number of slots in each frame exceeds 1000, and ratio of time-sensitive parts to all traffic is below 0.6, the calling loss ratio for a one-slot channel application is near to 0.

The utilization is variable according to the relationship of preserved bandwidth and bandwidth of each slot. When each appointed channel is full, the utilization is equal to length of type of slot bits / length of each slot=(512-2)/512=99.6%.

6.2 Filler Channel Transfer

The filler channels are used to transmit non time-sensitive traffics, in spite of its delay and jitter is increased, but these are permitted for non time-sensitive traffic.

It avoids the heavy congestion in the network, because of using accessing control and path preservation like MPLS.

Under the condition of light congestion, since MPsLS equips a long queue to store data when burst taking place, data loss ratio is kept very small. The calling loss depends on the relationship between bottleneck bandwidth in the cloud and peak bit rate value of arrived traffics. As type of slot bit occupies less proportion to whole slot, we only consider overhead caused by the non time-sensitive traffics. The utilization is minimum (512-32-2)/512=93.1% for a shot packet and maximum (512-2)/512 = 97.7% for a long packet.

7. Conclusions

MPsLS is a novel data forwarding scheme for the core backbone of Internet. It can provide complete QoS guarantee to the time-sensitive traffics in integrated network, at the same time, it accomplish good channel utilization due to the introducing of filler channel transmission.

MPsLS is also an extension of DTM technology. It provides both appointed channels to transmit synchronously time-sensitive traffics and filler channels to transmit non time-sensitive traffics in packet mode simultaneously. MPsLS has not only ability of QoS guarantee but also possibility of high channel utilization accomplishment.

In addition, when the ratio of time-sensitive traffics to total traffics is less than 60%, new calling loss is nearly equal to zero.

On the other hand, the utilization of network resources is greater than or equal to the utilization in ATM transmission.

However, for CBR and VBR applications, the appointed channels are preserved at peak bit rate. This leads to larger ratio of real-time traffics than really required, and result in decreasing of resources utilization. These wasted slots are used dynamically by filler channels. Therefore it results in high resource utilization.

References:

[1] C. Bohm, M. Hidell and P. Lindgren, "DTM—A True Broadband Concept Based on Fast Circuit Switching", In Proc. 6th ERCIM Workshop, Stockholm, Jun. 1994.

[2] C. Bohm, et al., "The TDM Gigabit Network", Journal of High Speed Network, Vol.3, No.2, 1994.

[3] L. Ramfelt, "Performance Analysis of Slot Management in TDM Networks", Technical Report TRITA-IT R 95:23, Dept. of Teleinformatics, KTH, Stockholm, Jan. 1996.

[4] B.Mikkelsen et al., "160 Gb/s TDM Transmission System", Proceeding of ECOC 2000, paper 6.1.1, Munchen, Sep. 2000.

[5] C. Bohm, et al., "The DTM Gigabit Network", http://www.cnaf.infn.it/~ferrari/papers/opt/93_Maryland_JHSN_DTM.pdf

[6]. L. Berger and Movaz, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[7]. P. Ashwood-Smith, et. al., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions", RFC 3472, January 2003.

[8]. L. Berger and Movaz, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.