Physical Frame Time-Slot Switching – A New Switching Technique over DWDM*

XU Deng-yuan¹, DOU Jun¹, TU Jian-ping²

(¹ School of Computer and Communications Engineering, Southwest Jiaotong University  Chengdu 610031  China; ² Jiangsu Automation Research Institute, Lianyungang 222006  Jiangsu  China)

Abstract  A novel technique called physical frame time-slot switching (PFTS) is discussed and its technical and application aspects are analyzed. The format of the ethernet media access control (MAC) frame is borrowed in defining the physical frame for PFTS and the transmission time for the maximum size of the MAC frame is defined as a physical frame time-slot (PFT). Consequently, user data can be fed into PFTs and switched in a single layer sub-network in an asynchronous mode.

Key words  physical frame time-slot switching; three dimension ethernet MAN; ethernet-like physical frame; single user plane architecture

The advent of dense wavelength division multiplexing (DWDM) raises the raw bit rate in a single fiber from a few Gigabits per second (Gbps) to tens of Terabits per second (Tbps), and enables the total bit rate in a single fiber reaching the order of magnitude of Tera bits. This achievement opens a new era for high-speed networks, and potentially, enables existing telephone, cable TV, and computer networks to be merged into a unified network.

However, in a merged network, service requirement for time critical traffic cannot be satisfied through statistically sharing a high-speed channel as is widely adopted today in Internet. With DWDM techniques, even the transmission capacity of a single lambda is much too big a channel for individual user-data transfer (10/20/40 Gbps at present, and 80 Gbps soon), therefore, a lambda channel needs to be further decomposed into sub-channels. Moreover, Quality of Service (QoS) for time-critical data flows requires a sub-channel with guaranteed throughput and other time-relevant parameters.

Existing techniques, such as optic burst switching (OBS), optic packet switching (OPS), etc., have been around for some time, where variable length of data bursts or packet in fixed length of time-slots are used as the basis for switching1[4]. These approaches utilize lambda switching technique to interconnect wavelengths among fiber ports in a switch, so as to establish a lambda path between ingress router and egress router. Lambda switching processes in all the switches along the path are controlled by burst head or packet head, which are transmitted through dedicated lambda in advance. In other words, they both adopt out-band signaling approach at the physical layer and data and control signals are transferred in different lambdas. Lack of mechanism in guaranteeing required QoS for time-critical traffic and low channel utility are main problems with these approaches.

This paper promotes a novel technique called the “physical frame time-slot switching (PFTS)” and a framework of PFTS is also presented. The format of the Ethernet MAC frame is borrowed in defining the physical frame for PFTS and the transmission time for the maximum size of the MAC frame is defined as a physical frame time-slot (PFT). Consequently, user data can be fed into PFTs and switched in a single layer sub-network in an asynchronous mode.

With out-band signaling concept, a PFTS sub-network in user plane (U-plane) together with existing Internet protocol stack in signaling and management planes (S&M planes) create a new architecture for Next Generation Internet, called the single user plane architecture (SUPA). The SUPA network with a Physical Frame Time-slot switching layer over DWDM can provide a virtual line switched service for time-critical traffic such as in cable TV and telephone switching service with guaranteed QoS.

1 Existing Switching Techniques over DWDM

1.1 Burst Switching

Burst switching or OBS does focus on relay of
individual data, called “bursts” – a representation of data with variable lengths from a few kilobytes up to few megabytes, hence a burst switching network handles user data in terms of big data-blocks[1,2].

Optical burst switching is an adaptation of an ITU (International Telecommunication Union) standard for burst switching in ATM networks, known as ATM block transfer (ABT). There are two versions of ABT: ABT with delayed transmission and ABT with immediate transmission.

People are in favor of pure optic approach in realization of burst switching in order to avoid the electric/optic conversion. Consequently, they have to try to push most processing and buffering functions from pure optic core-nodes to the photoelectric edge-routers in a burst switching network because of difficulty in optic buffering and lack of optic processing capability.

Burst switching adopts out-band signaling concept, where user data and control data are transmitted in different lambdas. This implies that burst switching has to rely on controllable lambda switching techniques within a node, otherwise it degenerated to pure fixed lambda switching.

Optic nodes have to be notified by the “burst head” (called “control packet” in OBS) through control lambdas so as to interconnect relevant wavelengths in advance; burst transmission will be postponed until a concatenated lambda path between edge routers at the optic burst domain boundary is reserved; otherwise, it has to bear the penalty of data loss. Alternatively, Fiber delay lines or delay loops can be used to control the delayed time in all the optic nodes along the data path while notification of burst is in forwarding.

Ilia Baldine, et al., gave up pure optic control approach in core nodes, and suggested a signaling approach called just enough time (JET), with which signaling “bursts” are done in an electric/optic combined manner[1].

The shortcomings of burst switching can be summarized as:

1) Lack of mechanism for guaranteed QoS. There is no mentioning of mechanism for ensuring user required QoS, especially for time-critical traffic where guaranteed throughput, transit delay, and other time parameters.

2) Complexity in parameter selection and adjustment. Switching of optic lambdas has to be done through electronically controlled optic lambda switches. Because of the difference in speed between optic transmission and electronically controlled lambda switching (including head processing), there is a need for user data buffering. However, for lack of buffering and processing capability in optic domain, different transmission strategies have to be suggested:

   1) TAG (Tell And Go) – a data burst is sent on the data lambda right after the control packet is sent out. This demands buffering of the burst in the next switch, which has to rely on fiber delay line (FDL), otherwise transmission contention and data loss may arise. Ref.[5] provides a figure which shows that a $10^4$ bits optic “buffer” needs a delay line of one kilometer fiber. Alternatively a fiber delay loop longer than the length holding a maximum burst is employed, where controlled times of circulations in the loop can multiply the delay time[6]. The difficulties with this approach lies in calculation and control of times of circulation.

   2) JET-Instead of sending data burst immediately in the ingress router, JET postpones data sending time called the “offset time” which is just enough for the establishment of lambda path, so as to avoid buffering in intermediate optic switches and possible data losses. With this approach, it would be difficult to determine the required offset time.

   3) Determining Length of Burst. Burst length and offset time (or buffering time) have a complex interrelationship. The longer the maximum length of the burst is, the higher the channel utility can be obtained. Consequently a longer offset time is needed. On the other hand, waiting for collection up to a maximum burst might result in longer waiting time at the ingress router. All these have made the tradeoff in determining parameters difficult.

Study of burst switching is still in its early theoretic phase, and lack of the processing power in the pure optic core node has already imposed many difficulties, although in the long run, pure optic network is very attractive. In view that the optic computing is still in its infancy, there is a long way to go before an optic burst switching node capable of handling hundreds of wavelengths, with acceptable price on the market.

1.2 Optic Packet Switching

Optic Packet Switching differs from burst switching chiefly by use a “packet” measured by a fixed duration[3]. One influential OPS project-KEOPS (KEys to optic packet switching) employs a 1.646 μs
time-slot including a 14-byte header and 1.35 μs for payload.

Generally speaking, there is not much difference between OBS and OPS in KEPOS, except that a burst length is variable while the length of an optic packet is fixed. Variable length might incur inefficiency in channel utility caused by fixed path establishment time, while the inefficiency in OPS occurs when user data is shorter than the fixed optic packet length.

The OPS has been successfully implemented in KEPOS with a low data rate control channel (622 Mbps) and an IP-oriented architecture has been exercised. Compared with OBS, it only works in low bit rate in a lambda. With out-band signaling approach in switching sub-layer, OPS inherits some of the shortcoming of OBS as discussed in section 2.1.

2 Physical Frame Time-slot Switching

Historically, there have been two opposite views with regard to services provision: “best effort” and “ensured”. The first view is represented by Internet service, where user data dynamically share a transport capability provided by the physical network, and the network service provider will try its best to meet the requirement according to network load and service policies. Even with the new strategies such as integrated service and differentiated service \[^7,8\]
control of QoS at IP layer or above cannot really guarantee the required QoS for lack of support mechanisms from lower layers.

An ATM based network can be viewed as taking the second view, where a user data flow is assigned to a virtual circuit (VC) of the data link layer according to its QoS requirement, and the VC is, in turn, bound to physical frames e.g. time-slots in SDH. Consequently, QoS parameters, such as throughput, transmission delay, etc., can be ensured by assigning proper numbers of time-slots to the data flow.

This section is focused on a novel technique called physical frame time-slot switching (PFTS), where the format of the Ethernet MAC frame is adopted as the Physical Frame (called Ethernet-like physical frame, EPF) and the transmission time for EPF is defined as the basic time-slot. PFTS works in an asynchronous mode without the need for global clock synchronization and can provide a virtual line-switched service for time-critical data flows. For lack of granular multiplexing mechanism within a lambda of DWDM, the DWDM architecture should be augmented to support multi-class of QoS requirement. 

2.1 Augmentation to Existing DWDM Layer

Existing DWDM devices can be conceptually viewed as having a two sub-layer structure: the MDTRS (Medium Dependent Transmission and Receiving Sub-layer) at the bottom and the LMDS (Lambda Multiplexing/De-multiplexing Sub-layer) on the top. The LMDS does multiplexing and de-multiplexing functions on the basis of lambdas, while MDTRS does transmission and receiving functions for multiplexed lambdas to or from a fiber.

Lambda switching cannot satisfy the general requirement in a multi-user network environment. To cater for this requirement, the architecture can be enhanced with a physical frame sub-layer on top of LMDS.

In view that the Ethernet is the most popular technique in LAN, MAN, and soon in WAN as well, the format of the Ethernet MAC frame is deliberately chosen for the physical frame, and consequently, the new sub-layer is referred to as EPFS (Ethernet-like physical frame sub-layer). Fig.1 depicts an augmented architecture for DWDM physical layer.

![Fig. 1 An augmented DWDM architecture](image)

People in favor of pure optic switching have argued that data handling in electronic domain is much slower than that in optic domain, and E/O conversion also takes time; therefore, E/O combined approach towards physical frame switching introduces excessive time delay and might be a source of jitters. The opposite argument would be: such a delay indeed will postpone the time of the arrival of the first frame, but it will not change the time-relationship between frames within a data stream, nor it will generate jitters as long as such delay and conversion time remain comparatively constant. Indeed, such argument can be backed up by the success in 10 GE products and 40 Gbps Chip \[^9,10\]. More importantly, it would be much
more comfortable to deal with buffering and necessary processing electronically than to do it optically considering the state of art in two domains.

With respect to E/O conversion, an experiment with super-conducting techniques has demonstrated that a conversion rate of 100 Gbps is reachable. To save the processing time for frame header, parallel techniques, optimal look-up algorithm for routing, and hardware based processing should be applied.

Fig. 1 also illustrates the relationship between the EPFS and its upper Ethernet MAC Sub-layer. Mapping between MAC frame and the EPF is straightforward since they have the same format except for a minor modification to the destination address field as explained in the next section.

### 2.2 Operation of Physical Frame Time-Slot Switching

The addition of EPFS on top of existing DWDM enables users to mingle their data into physical bit streams in the form of formatted frames. Here, a novel concept called the “physical frame time-slot” (PFT) plays an important role in multiplexing within a lambda, in relaying frames within a DWDM node, and in synchronization between DWDM nodes.

A PFT is defined as the transmission time for an EPF with the maximum MAC frame length (1542 bytes) defined by IEEE 802.3 (2002 version), and used as the basis for frame-based asynchronous operation. An ingress edge router of PFTS domain will simply to hand received MAC frames down to the physical layer, which will, in turn, treat them as EPFs and multiplex them into the bit stream of different lambdas towards PFTS nodes according to routing decision. DWDM switching nodes will then forward PFTs on the DWDM physical layer network along the predetermined path to egress edge router and synchronize their operation by detecting the frame preambles. This enables an asynchronous operation based on PFTs without a need for clock synchronization among switching nodes.

The format of an EPF resembles an Ethernet MAC frame except that the destination address field in the MAC frame is used in the EPF as the switching field, which consists of a PN (port number), an LN (lambda number), and a VC (virtual circuit) Id. Figures 2 and 3 illustrate the switching model for PFTS nodes and the format of switching field in an EPF frame respectively.

![PFT switching node model](image)

Fig. 2 is essentially two sets of sub-layer stratum in Fig. 1 interconnected by a Physical Frame Time-slot (PFT) relay entity, which switches individual PFTs according to the fields of the Physical Frame shown in Fig. 3.

As shown in Fig. 3, the PN is used in DWDM nodes to identify a physical port (corresponding to a fiber), while the LN uniquely identifies a single lambda within a given fiber. VC Identifier is designed to distinguish different user data streams. A VC is also associated with certain transmission capability decided by the number of PFT(s) per second assigned to that VC. Finally, the Router Id is reserved for shared media topologies, such as a ring, where a simple number is used to replace complicated MAC address.

![Reassignment of the destination address](image)

Fig. 3 Reassignment of the destination address

### 2.3 Advantages of PFTS

PFTS differs from the existing switching techniques such as OBS, OPS in many ways, and its advantages can be summarized as follows.

#### 2.3.1 High Channel Utility

First of all, its theoretical channel utility can approach 100% as long as all the time-slots are used in carrying user traffic, since PFTS is time-slotted. Furthermore, the theoretical effective data rate can be as high as 97.3% with the maximum extended frame length (1542 bytes) [IEEE 802.3 2002] against the maximum payload length (1500 bytes), which are much superior over other switching techniques discussed in section 1.
2.3.2 Capable of Provision of Virtual Line-Switched Service with Guaranteed QoS

PFTS service is connection-oriented and most suitable continuous traffic flows, which are dominated traffic in B/S application model in Internet today and will form an even greater percentage in the traffic of future merged networks (cable TV streams, switched telephone traffic, and other audio/video streams). Communication path or virtual physical layer connection between an ingress router and egress router needs to be set up once for one communication session, which simplifies the routing process into “IP routing once, and fast switching for whole session”. Compared with manner of “setting up a path for each burst or packet” in OBS and OPS, it is much more efficient.

QoS provided by PFTS can be guaranteed when enough time-slots are assigned to a virtual physical layer connection with fast transmission in the fiber and fixed processing time in each node. Techniques such as RSVP and CR-LDP (constraint-based routing label distribution protocol) developed for Internet or MPLS, can all be utilized.

2.3.3 Merging the Data Link and Physical Layer into a Single Layer

The PFTS enables a user plane to be reduced to a single physical layer, in which some of the data link layer functions such as framing, error checking, and etc., can be embedded. As the Ethernet has been widely adopted not only in end-systems but also in intermediate systems (sub-networks), MAC frame used as physical layer frame in PFTS can save lots of troubles of segmentation and reassembly between various data link frames. It has been envisaged that the data link layer and physical layer will eventually be merge into a single layer, not only in the user plane but also in signaling and management planes.

3 Single-layer User Plane Architecture for Next Generation Internet

Introduction of PFTS technique is part of the research result in the 3D-EMAN project carried out at Sichuan Network Communication Key Laboratory. The 3D-EMAN architecture utilizes current Internet protocol stack in C & M planes and the PFTS on top of DWDM in U-plane[12]. Operation between systems supporting 3D-EMAN architecture will operate in 3D-EMAN mode, where Internet protocol stack plus a moderate enhancement will play the C & M functions and transport user data on a PFTS network. For SVC service, a path between ingress and egress edge routers can be established by use of Internet routing protocols and QoS negotiation is done through 3D-EMAN specific protocol. PVC service, such as cable TV broadcasting service, can be subscribed in advance. Detailed discussion concerning 3D-EMAN can be found in Ref.[12].

PFTS applied to 3D-EMAN can create a new architecture called Single-layer user plane architecture (SUPA). The SUPA network with a PFTS sub-layer over DWDM can provide a virtual line switched service for time-critical traffic such as in cable TV and telephone switching service with guaranteed QoS.

Through the experience gained with 3D-EMAN at Sichuan Network Communication Key Laboratory, the authors are convinced that SUPA with the PFTS for next generation Internet is feasible.

Acknowledgement

The work presented in this paper has been sponsored by Ph.D student Creative Foundation of Southwest Jiaotong University. The authors would like to acknowledge the financial support from it.

References


(Continued on page 263)