New framework and performance evaluation for IP DiffServ over OBS networks[*]

LONG Ke-ping[1], Rodney S. Tucker[2], WANG Chong-gang[3]

(1. *Special Research Centre for Optical Internet &. Wireless Information Networks,*
Chongqing University of Posts &. Telecommunications, Chongqing 400065, *P. R. China;*
2. *ARC Special Research Centre for Ultra Broadband Information Networks, University of Melbourne,*
Victoria 3010, *Australia;* 3. *HongKong University of Science and Technology, HongKong, P. R. China*)

Abstract: IP Differentiated Services (DiffServ) has been standardized by the IETF and is considered as a promising IP QoS solution due to its scalability and ease of implementation. In this paper, we present a novel framework for IP DiffServ over optical burst switching (OBS), namely, DS-OBS. We present the network architecture, functional model of edge nodes and core nodes, the control packet format, a novel burst assembly scheme at ingress nodes and scheduling algorithm of core nodes. The basic idea is to apply DiffServ capable burst assembly at ingress nodes and perform different per hop behavior (PHB) electronic treatment for control packets of different QoS classes service at core nodes. Simulation results show that the proposed schemes can provide the best differentiated service for expedited forwarding (EF), Assured forwarding (AF) and best effort (BE) service in terms of end-to-end delay, throughput and IP packet loss probability.

Key words: IP QoS, differentiated services; per hop behavior; optical burst switching; DS-OBS

CLC number: TN929. 1　　**Document code:** A

0　Introduction

IP over WDM has been envisioned as the dominant network architecture for the next generation optical Internet since it eliminates the intermediate layers and can make better use of advanced optical technologies[1,2]. Currently, optical burst switching (OBS) is under study as a promising switching paradigm for the optical backbone network in IP over WDM networks[3-5]. The basic idea of OBS is to decouple the data channel from the control channel and to combine the best of the coarse-grained optical circuit switching and the fine-grained optical packet switching in order to facilitate the efficient integration of IP and WDM.

Due to the increasing deployment of Internet applications requiring QoS, such as voice and video application, a challenge issue for IP over WDM networks is how to support IP QoS at the WDM layer. Among the different IP QoS solutions, the Differentiated Services (DiffServ) architecture[6] is currently considered as a promising IP QoS solution due to its scalability and ease of implementation in electronic switching and routing techniques. An end-to-end DiffServ is achieved by concatenation of per-domain services and service level agreements (SLAs) between adjoining domains along the route that the traffic crosses from the source to the destination. The per-domain service is realized by providing qualitative per hop behavior (PHB) using active queue management and scheduling algorithms at the core router for packets with different marking at DiffServ codepoints (DSCP) marked by the ingress router. Two PHBs are also standardized by the IETF, there are: Expedited Forwarding PHB(EF PHB)[7] to implement services requiring low loss and low delay; and assured forwarding PHB(AF PHBs)[8] to implement services requiring assured bandwidth. However, these buffer-based schemes are difficult to apply to the WDM layer because optical RAM is not yet available even though a limited delay can be provided to optical packets by

fiber delay lines (FDLs) .

Recently, an offset-time-based QoS scheme for optical burst switching networks has been proposed by Qiao[9] to provide basic QoS by only considering the metric of burst blocking probability (BBP). Using this scheme, Qiao has analyzed the degree of class isolation using the offset times, upper/lower bounds on burst blocking probability, and QoS performance in terms of burst blocking probability and queuing delay[9]. Dolzer[10] has extended model of the offset-time-based QoS scheme by considering the interaction of offered load between the higher priority traffic and lower priority traffic. Their analysis and simulation results show that the offset-time-based QoS scheme can achieve lower burst blocking probability for higher priority traffic while degrading the performance of lower priority traffic.

However, there are several key issues are unclear for the offset-time-based QoS scheme. First, there have not been any reported studies on how the offset-time-based OBS QoS scheme supports the DiffServ or IntServ standards of the IETF. Secondly, the end-to-end (at least edge-to-edge) IP Packet QoS performance for the offset-time-based QoS scheme does not appear to have been analyzed or evaluated. The QoS performance of bursts is different to the QoS performance of IP packets. Chen[11] argues that the end-to-end delay for data bursts will not represent the end-to-end delay of IP packets since packets of a real-time application may be separated by different data bursts at the burst assembly stage. Third, all existing studies assume that no contention occurs between control packets (reservation requests). As discussed in[12], this assumption is incorrect when multiple control packets (reservation requests) from different edge routers arrive at a core node synchronously, especially in the case of a fixed timer-based burst assembly algorithm and constant offset-time setup.

In this paper, we address the above key issues and proposed a novel framework for IP DiffServ over OBS networks, namely, DS-OBS. This paper is organized as follows. In section 1, we define a new format for burst control packets and present a DiffServ capable OBS network architecture and functional model of edge/core nodes. In section 2, we propose a novel DiffServ-capable burst assembly schemeat ingress nodes. Extensive simulation results in terms of end-to-end delay, throughput and IP packet loss probability for

proposed architecture and schemes are given in section 3. Section 4 concludes this paper.

1　Network architecture and node model for IP DiffServ over OBS

1.1　Network architecture for IP DiffServ over OBS

The proposed network architecture and the basic function of edge routers and core routers for IP DiffServ over OBS (DS-OBS) are shown in fig. 1, which consists of ingress nodes, core nodes and egress nodes. The basic idea of the proposed architecture is to provide IP DS-OBS while avoiding the deployment of extra offset and optical buffers. In contrast to the traditional OBS network architecture. In our proposed architecture, when
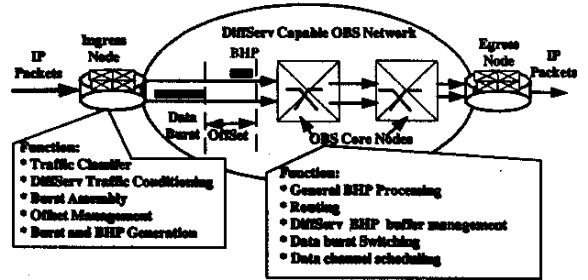

Fig. 1 **Network architecture for IP DiffServ over OBS**

traditional IP packets enter the ingress router that integrates the function of a DiffServ edge router with the function of an OBS edge router, the functions of both traffic conditioning and burst assembly are performed for the IP packets. Then, a data burst and a proposed control packet carrying the DSCP information as well as other information are generated and transmitted into the core of the OBS network separately on different wavelengths with the Burst Head Packets (BHPs) sent ahead by an offset time. At core nodes, BHPs are processed electronically to provide different per-hop behaviors (PHBs) for different classes of service. As a result of this processing, the differentiated service for corresponding data bursts was implemented. At the egress node of a domain, the data bursts are disassembled back into IP packets to be forwarded to their next hops (e. g. traditional IP routers).

1.2　Functional model of ingress nodes and control packet

The proposed functional model of an ingress router and the control packet format for the DS-OBS are shown in fig. 2. After being marked as EF, AF and BE service, the IP packets are enqueued into separate queues. The burst assembler performs DiffServ-capable assembly for

EF，AF and BE service by adaptively adjusting theassembly parameters（timer and burst size） based on the QoS requirements of different classes of service and the metering results of actual traffic arrivals. A detailed burst assembly algorithm will be given in section 2.
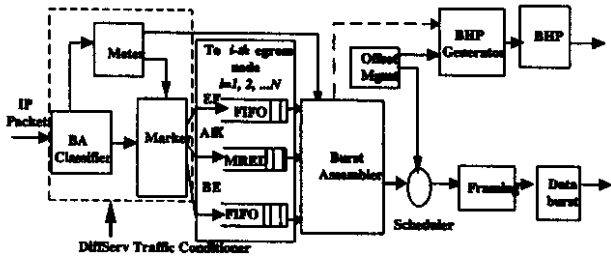

Fig. 2 **Model of ingress node and control packet**
format for DS-OBS network

Once a burst of EF，AF or BE service has been created， based on the information of burst assembly and offset management， control packet generator will generate a burst header packet （BHP）， which is defined as fig. 2. This includes information such as DSCP， offset time， source/ destination address and burst size. From fig. 2，we can see the maximum length of a BHP is less than 45 bytes and 128 bits of Source/Destination add-ress is reserved for the future IPv6 protocol. So far，no standardization of OBS control packets has been reported in the existing literatures. We define a new BHP not only for our scheme but also for being extended to other OBS networks， e. g. if （G）MPLS is deployed for traffic engineering in the OBS network，we can only change the of source/ destination address fields to a "Label" field with 20bits length.

In this paper， both the constant offset and random offset settings are considered. As discussed in ［12］， in order to prevent excessive loss due to the synchronized arrivals of data bursts at the same output of OBS core nodes， a random offset generation (e. g. Possion token bucket) is deployed for data burst shaping.

1. 3　Functional model and scheduling algorithm of core nodes

The control part of a core router is shown in fig. 3，which keeps the optical data processing as simple as possible. The priority treatment of data bursts is performed by the different per-hop behavior treatment for their BHPs of EF， AF and BE services. This is achieved by priority
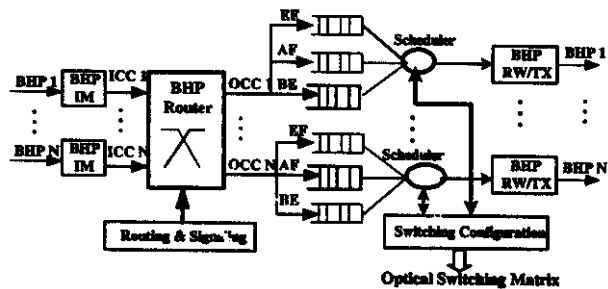

Fig. 3 **Control part of a core node for DS-OBS network**

scheduling for BHPs and appropriately configuring the transmission rate of control channel as follows.

$$Num_{hop} \times Num_{datachannel} \times \frac{Avg_{BHPSize}}{Rate_{controlchannel}} = \frac{Avg_{burstsize}}{Rate_{datachannel}}$$

(1)

In contrast to existing schedulers in OBS core nodes， the scheduler presented in this paper as shown in fig. 3 has two functions. First， it performs priority scheduling （or Weighted Round Robin scheduling， WRR） for multiple BHPs queues of EF， AF and BE service according to the information carried in the BHPs. As a result of differentiated service for BHPs of different service classes， the differentiated service for the corresponding data bursts was achieved. Secondly， it schedules data bursts on outgoing data channels using data channel scheduling algorithms， e. g. the latest available unused channel with void filling （LAUC-VF） algorithm[13] and our proposed the Least Gap with Void Filling（LGVF） algorithm[15]. Since the main differentiated treatment has already been conducted for the BHPs of different service classes， the performance enhancement of our LGVF over LAUC-VF in the DS-OBS architecture is not obvious as the enhancement in other OBS network[15].

The scheduler is bidirectionally connected to the switching configuration module. The aim is to exchange configuration information between them， to make the scheduler update the state information of control channel and data channels， to rewrite fields of BHPs and transmit the BHPs to the next hop.

2　DiffServ-capable burst assembly scheme

Several burst assembly algorithms have been proposed in ［13，16］. Xiong[13] proposed a bucket-based burst assembly scheme based on the egress router addresses， assembly time intervals and maximum data burst size. There are several open issues in this scheme， such as：how to dimension the parameters of buckets according to the traffic

load and number of QoS classes and destinations，burst assembly time T_c　and maximum burst size L. To achieve a better trade-off between the burst assembly delay and burst assembly efficiency when input traffic load is temporarily low，An Ge[16] presented a simple burst assembly mechanism based on a time-window and minimum data burst size. However，both of these two mechanisms cannot support IP DiffServ and cannot adjust the parameters of burst assembly according to the QoS characteristic of different service classes.

In this section，we propose a novel DiffServ-based burst assembly scheme. The proposed scheme can support the DiffServ services（EF PHB，AF PHB and BE）standardized by the IETF，and can adjust the parameters adaptively according to the actual traffic arrival rate（output of meter as shown in fig. 2）and the differentiated QoS requirements of different service classes. The detailed burst assembly algorithm is given as follows.

Definition：

T_A^{EF}, T_A^{AF}: Maximum burst assembly time of EF and AF service，configured by the delay requirement of EF and AF.

T_A: Allowable maximum burst assembly time in an OBS network. Dimensioned based on the number of destinations，the number of QoS classes and the end-to-end delay.

MAX_{BS}, MIN_{BS}: Allowable maximum/minimum burst size in an OBS network. Decided by the allowable burst assembly delay，burst assembly efficiency and link utilization[12,13].

$BS_{min}^{EF}, BS_{min}^{AF}, BS_{min}^{BE}$: Allowable minimum burst size of EF，AF and BE service. BS_{min}^{AF} is configured to be more than MIN_{BS} and is decided by the Committed Information Rate（CIR）in the metering using trTCM [14]. BS_{min}^{BE} is configured to be equal to MAX_{BS}.

$BS_{max}^{EF}, BS_{max}^{AF}, BS_{max}^{BE}$: Allowable maximum burst size of EF，AF and BE service. Considering the delay characteristic of EF PHB service，set the $BS_{min}^{EF} = BS_{max}^{EF} = MIN_{BS}$. BS_{max}^{AF} is less than MAX_{BS} and is decided by the Peak Information Rate（PIR）in the metering using trTCM[14].

$T_k^{EF}, T_k^{AF}, T_k^{BE}$ is the timer for assembling EF，AF，BE，respectively.

$L_{EF}(k), L_{AF}(k), L_{BE}(k)$ is the assembly meter for computing the assembled burst length of EF，AF，BE，respectively.

Algorithm：

① When a packet with length of pkt_size arrives at a queue associated with a destination at the ingress node of the OBS domain.

If it is marked as EF PHB and $L_{EF}(k) = 0$，start timer T_k^{EF} and $L_{EF}(k) = L_{EF}(k) +$ pkt_size.

Else if it is marked as AF PHB and $L_{AF}(k) = 0$，start timer T_k^{AF} and $L_{AF}(k) = L_{AF}(k) +$ pkt_size.

Else it is marked as BE and $L_{BE}(k) = 0$，start timer T_k^{BE} and $L_{BE}(k) = L_{BE}(k) +$ pkt_size.

② If$\{T_k^{EF} > T_A^{EF}$ or $L_{EF}(k) > BS_{min}^{EF} = MIN_{BS}\}$，the EF burst is created and reported with the length of burst $L_{EF}(k)$ or MIN_{BS} if $L_{EF}(k)$ is less than MIN_{BS} when the timer T_k^{EF} reaches the T_A^{EF}，reset $T_k^{EF} = 0$ and $L_{EF}(k) = 0$.

Else if $\{T_k^{AF} > T_A^{AF}$ or $BS_{min}^{EF} < L_{AF}(k) < BS_{max}^{AF}\}$，the AF burst is created and is reported with the length of burst $L_{AF}(k)$ or BS_{min}^{AF} if $L_{AF}(k)$ is less than BS_{min}^{AF} when the timer T_k^{AF}，reaches T_A^{AF}，reset $T_k^{AF} = 0$ and $L_{AF}(k) = 0$.

Else if $\{T_k^{BE} > T_A$ or $L_{BE}(k) \geq MAX_{BS}\}$，the BE burst is created and reported with the length of burst $L_{BE}(k)$ or MAX_{BS} if $L_{BE}(k)$ is less than MAX_{BS} when the timer T_k^{BE} rea the T_A，reset $T_k^{BE} = 0$ and $L_{BE}(k) = 0$.

③ Otherwise，keep all the values unchanged and waiting for the next IP packet arrival and go to ①.

3　Performance evaluation

Using OPNET，we have developed simulation models of our proposed architecture，protocol，burst assembly scheme at ingress nodes and scheduling algorithm at core nodes. We have simulated the end-to-end IP performance and IP packet loss rate at core nodes using the simulation network shown in fig. 4. Each of four Possion traffic sources（one EF，one AF and two BE）with the same traffic arrival rate and same average packet length（500 bytes）was accessed by each of four ingress routers via 2. 5 Gbit/s wavelengths and were multiplexed into core router 0. The multiplexed stream was transmitted to each of four egress nodes via core routers over 4-hops with $4 \times$ 2. 5 Gbit/s data wavelengths. The BHPs are transmitted on a separate channel in the core of network.
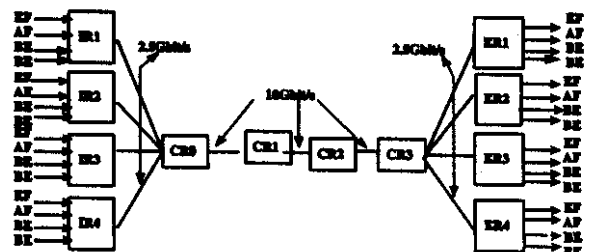

Fig. 4 Simulation network topology

Three scenarios are simulated to evaluate the

performance of our schemes: scenario 1, constant offset at ingress nodes, FIFO for BHPs at core nodes; scenario 2, constant offset at ingress nodes, WRR at core nodes; scenario 3, random offset at ingress nodes, WRR at Core nodes. The parameters of the burst assembler are shown in Table 1.

Table. 1 Parameters of burst assembler

Service classes	Min burst size (kBytes)	Max burst size (kBytes)	Timer (ms)
EF	5	5	4.8
AF	30	50	55
BE	125	125	600

Fig. 5 shows the IP packet loss rate at Core0 as a function of the source traffic load, which was caused by both the congestion of control packets and the scheduling failure of data bursts. The IP packet loss rate in all three scenarios is low due to the proposed burst assembly scheme at ingress nodes. As the load increases, the packet loss rate increases. However, The IP packet loss rate is acceptable even if the load is 1.0. Moreover, the best performance is achieved by using random offset setting at ingress nodes and WRR at core nodes. When a constant offset is applied, the performance of WRR scheduling is better than the performance of FIFO.
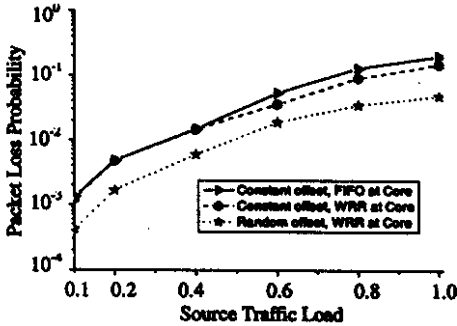

Fig. 5 IP packet loss rate at Core0 in the three scenarios

Fig. 6 shows the actual offered load at the input of Core0 and the link utilization at the output of Core0 as a function of source traffic load. The link utilization of all three scenarios is high due to the low IP packet loss rate. The link utilization for the scenario with random offset at ingress nodes and WRR at core nodes is just slightly lower than the actual offered load at Core0 due to the very low packet loss rate even if the source traffic load is 1.0. This is because DifffServ traffic conditioning (metering, marking, shaping, dropping) and admission control have been applied for bursty IP traffic to assure their performance and to avoid the waste of resources at the core of the DS-OBS network. So, the actual offered load at input of Core0 is less than source traffic load.
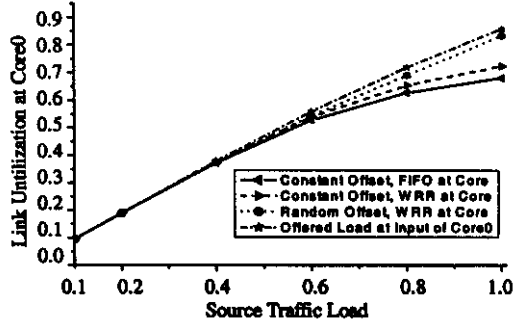

Fig. 6 Link utilization at Core0

Furthermore, in order to illustrate the service differentiation of our proposed architecture, the packet loss rate for EF, AF and BE service under all three scenarios when the traffic load is 1.0 is shown in fig. 7. The scenario with random offset at ingress nodes and WRR at core nodes can achieve the best service differentiation and achieve the lowest packet loss rate for AF and EF as well as the lowest total packet loss rate. Scenario 2 (constant offset at ingress nodes, WRR at core nodes) has achieved better service differentiation and total packet loss rate than scenario 1 (constant offset at ingress nodes, FIFO at core nodes).
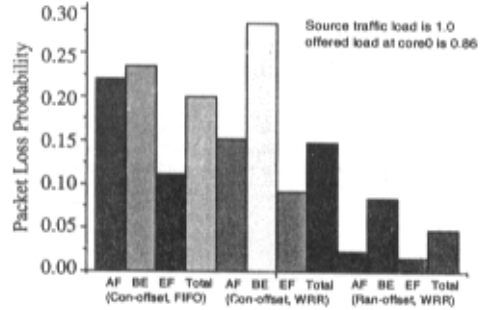

Fig. 7 The packet loss rate for EF, AF, BE in three scenarios (load 1.0)

The actual achieved end-to-end (E2E) throughput normalized by reserved bandwidth for EF, AF and BE service as a function of source traffic load is given in fig. 8. When the load is below 0.4, three types of service can achieve over 0.92 times the expected bandwidth and there is no obvious difference among the three service classes. As load increases from 0.6, the service differentiation becomes obvious following the idea of DiffServ architecture for all three scenarios: AF service achieved the most assured bandwidth and EF service achieved as much bandwidth as they can due to the strict delay requirement. The bandwidth assurance of AF and EF service is achieved by decreasing the end-to-end throughput of BE service. The performance of setting random offset at ingress nodes and WRR at core nodes is the best compared to the other two scenarios. When the

constant offset is applied, the performance of WRR scheduling is better than FIFO. Results show that our DS-OBS is very efficient to support the IP DiffServ architecture.
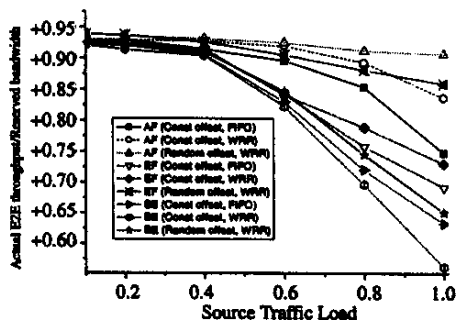

Fig. 8 Achieved end-to-end throughput normalized by reserved bandwidth

For each of AF, EF and BE service, E2E delay under the three scenarios is almost the same and is shown in fig. 9. Due to our Diffserv-capable burst assembly and since no extra offset is deployed for QoS, E2E Delay of EF service is less than 6. 67 ms even if the load is 0. 1. So, the possible worst E2E delay of EF service in 100-hops is about 160 ms and meets the delay demand of voice and video service. However, the E2E delay of AF is tens of milliseconds and E2E delay of Best effort (BE) is hundreds of milliseconds.
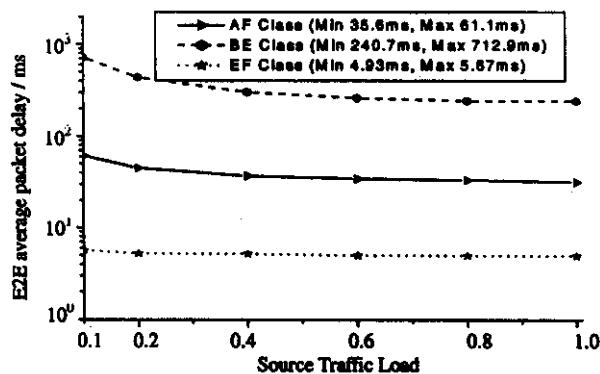

Fig. 9 End-to-end delay for EF, AF and BE for 4-hops network

4 Conclusion

In this paper, we propose a novel framework for IP DiffServ over OBS, namely DS-OBS, including the network architecture, novel functional model of ingress nodes and core nodes, a control packet structure, a DiffServ-capable burst assembly scheme at ingress nodes, and scheduling scheme at core nodes. Using a 4-hop network, we have simulated IP packet loss probability at core nodes, and E2E throughput and E2E delay for EF, AF and BE service. The results show that the proposed schemes can not only provide the best

service differentiation for EF, AF and BE in terms of both the E2E performance and IP packet loss probability at core nodes, but also enhance the whole performance of network.

References:

[1] BONENFANT P, RODRINGUEZ-Moral A and MANCHESTER J. IP over WDM:the missing Link [S]. White Paper, Lucent Technologies, 1999.

[2] GNHANI N, DIXIT S and WANG T. On IP-over-WDM integration [J]. IEEE Communication Magzine, 2000, 38(3):72-84.

[3] QIAO C and YOO M. Optical burst switching (OBS)—a new paradigm for an optical internet[J]. High Speed Networks, 1999, 8(1):69-84.

[4] QIAO C. Labeled optical burst switching for IP-over-WDM Integration [J]. IEEE Communication Magazine, 2000, 38(9):104-114.

[5] TURNER J. Terabit burst switching[J]. High Speed Networks, 1999, 8(1):3-16.

[6] BLAKE S, BLACK D, CARLSON M. An architecture for differentiated services [S]. IETF RFC 2475, Oct. 1998.

[7] JACOBSON V, NICHOLS K and PODURI K. An Expedited Forwarding PHB[S]. IETF RFC 2598, June 1999.

[8] HEINANEN J, BAKER F, WEISS W, et al. Assured forwarding PHB group[S]. IETF RFC 2597, June 1999.

[9] YOO M, QIAO C and SUDHIR Dixit. QoS performance of optical burst Switching in IP-over-WDM networks [J]. IEEE J. Select. Areas Commun. 2000, 18(10):2062-2070.

[10] DOLZER K, GAUGER C, SPITH J, et al. Evaluation of reservation mechanisms for optical burst switching [J]. AEU International Journal of ElectronicsandCommunications, 2001, 55(1):62-68.

[11] CHEN Y, HAMDI M, TSANG D, et al. Proportional QoS over OBS networks [C]. Proc. IEEE Global Telecommunications Conference (Globecom' 01), San Antonio, Nov 2001.

[12] CHASKAR H M, VERMA S and RAVIKANTH R. Optical burst switching:a viable solution for tarabit IPbackbone[J]. IEEENetwork, 2000, 14(6):48-53.

[13] XIONG Yi-jun, MARC Vandenhoute. Control architecture in optical burst-switched WDM networks [J]. IEEE J. Select. Areas Commun. 2000, 18(1):1838-1851.

[14] HEINANEN J, GUERIN R. A two rate three color marker[S]. IETF RFC 2698, Sept. 1999.

[15] LONG Ke-ping, TUCHER R S, OHSeen-Yoon OH. Fairness scheduling algorithms for supporting QoS in optical burst switching networks[C]. Proceedings of the SPIE's International Sympo sium: Asia-Pacific Optical and Wireless Communications (APOC), Shanghai, CHINA, 2002.

[16] AN Ge, FRANCO Callegati, LANKSHMAN S. Tamil. On optical burst switching and self-similar traffic [J]. IEEE Communication Letters, 2000, 4 (3):98-100.

theoretical reference for the analysis of capacitors with different structures，furthermore，it shows a principle for the accuracy estimation for the computation of capacitance of capacitors in engineering，thus it has some values in various applications.

References：

［1］ CHEN De-xiang. Discussion on distribution of external forces，charge and field in pulling a copper Plate out of a Capacitor［J］. College Physics，2002，21(11)：18-20.

［2］ GURU B S，HIZIROGLU H R. Electromagnetic fields and electromagnetics ［M］. Translated. by ZHOU Ke-ding，Beijing：Mechanic Industry Press，2000.

［3］ KRAUS J D，FLEISCH D A. Electromagnetics with applications ［M］. Beijing：Tsinghua University Press，2001.

［4］ XIE Chu-tang，RAO Ke-jin. Electromagnetic fields and electromagnetics ［M］. Beijing：Higher Education Press，1987.

(:   )

＊

，

(，   ，400065)

:   ，
，   ，
。

:   ；   ；

∿∿

IP DiffServ over OBS

＊

[1]，**Rodney S. Tucker**[2]，   [3]

(1.   ，   400065；
2.   ，   3010；3.   )

:IP DiffServ   IETF   ，   IP QoS   。
                IP DiffServ   ，   DS-OBS，   、
            、   、   。   offset time   OBS QoS
        ，   :   ，
                (PHB)   ，   。   :   、   IP
            EF   、AF   BE   。

:IP   ；   ；   ；   ；