

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

**ANSI***net*

Asian Network for Scientific Information  
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

## Research on SCTP-Based TML for ForCES Routers

Bin Zhuge, Huanhuan Song, Cheng Yu and Wei-ming Wang  
College of Information and Electronic Engineering, Zhejiang Gongshang University,  
Hangzhou, 310018, P.R. China

---

**Abstract:** The communication bottleneck between Control Elements (CEs) and Forwarding Elements (FEs) of ForCES router is the critical factor that restricts its large-scale application. In view of this problem, this study is about the communication efficiency and transmission performance of Transport Mapping Layer (TML) based on SCTP. It analyzes the coordination mechanism of channel flows among FEs. This study puts forward a complete flow control mechanism in internal channel of ForCES router and improves the communication efficiency. Furthermore, according to the distributive rate, the dispersion count and the flow rate of task, this study analyzes the performance of model. It shows that the flow control mechanism this study suggests can effectively guarantee throughput and reliability of communication between CEs and FEs. At last, this study verifies the accuracy of the theory through ns-2.34 simulation experiments.

**Key words:** Flow control, ForCES, transport mapping layer, bandwidth allocation, scheduling mechanism, flow prediction

---

### INTRODUCTION

The increasing Internet scale puts forward higher requirements to processing power, message transmission performance and physical interface support density of the network equipment. With the development trend of technology, the products of the next generation network had increasingly become modular, distributed, open and programmable (Fu, 2008). The new network architecture realized the modularization of network elements and open and unified interfaces were taken in the communication among the modules (Abed *et al.*, 2011). The new architecture could improve the scalability, flexibility and effectiveness (Fu, 2008). Forwarding and Control Element Separation (ForCES) of IETF with the above characteristics was the new generation network architecture (Doria *et al.*, 2010). The ForCES protocol which divides network elements into control elements and forwarding elements is used to achieve the cooperation and interaction among components. Thereby the scalability and manageability of network system were increased and the scalability and flexibility of network were enhanced (Takourout *et al.*, 2006; Fu, 2008).

In early 2004, General Router Management Protocol (GRMP) (Wang and Dong, 2004), Forwarding and Control Element protocol (FACT) (Audu *et al.*, 2003) and Netlink2 (Salim *et al.*, 2003) of candidate protocols were amalgamated into single one by the ForCES working group and it was ForCES protocol (Fu, 2008);

Khosravi and Anderson, 2003; Yang *et al.*, 2004). The ForCES working group respectively completed the ForCES requirements document (RFC 3654) and frame document (RFC 3746) in 2003 and 2004. Since then, the research had focused on the ForCES protocol (Halpern, 2005), FE model, LFB definition Library, ForCES TML, ForCES MIB and the design of standard draft document. In early 2010, four core drafts of the ForCES working group became RFC. It promoted manufacturers and scholars to have more intensive study on the architecture of ForCES network (Audu *et al.*, 2003; Halpern and Ma, 2007).

This study studies the flow control technology of the internal channels of ForCES router which is mainly based on the ForCES Transport Mapping Layer (TML). TML was important component of ForCES architecture network (Wang *et al.*, 2004) and also the key in restricting communication performance of the entire system. TML separates the protocol layer from the bottom layer of network. That solves the transmission problems of ForCES protocol messages in different link layers (such as, TCP/IP, ATM, Ethernet layer, backboard bus and so on). Thus, it provided a transparent, safe transmission channel for the ForCES protocol layer and improved the flexibility of ForCES system (Feng *et al.*, 1999; Aweya, 2001).

More and more units joined the ForCES team at home and abroad, because ForCES protocols officially became RFC international standards. The present study was based on the function of ForCES network elements and

the theory study on TML of ForCES network was relatively weak (Ogawa *et al.*, 2011; Dong *et al.*, 2011). This study is mainly about theoretical analysis and algorithm research on realization of some critical technologies of ForCES TML, certificated by software simulation and system experiment. The results will effectively promote the application development of protocol in ForCES TML and will make a solid foundation for the application of network element (Crouch *et al.*, 2010; Zhuge *et al.*, 2011; Mohammadi *et al.*, 2009).

**STRUCTURE OVERVIEW FOR FORCES TML COMMUNICATION MODEL**

ForCES TML communication model structure is the foundation to research the flow control technology in internal channels of ForCES router. The performance of communication model is directly related to the properties of the entire ForCES router. The separation of forwarding elements and control elements was the core of ForCES architecture router. RFC 3746 (Yang *et al.*, 2004), the ForCES framework document, defined the basic structure of a standard ForCES architecture network as Fig. 1 (Zhang *et al.*, 2009; Ibrahim, 2008).

In the ForCES architecture network, there is one or more CEs, they are main control CE and redundant CEs. Redundant CEs are mainly used to achieve the high availability of system. There can be many FEs in the network element and all of them are separated from the CEs. The interface between CE and FE is standardized, and the messages interaction conforms to ForCES protocol. The ForCES network element looks like a network node upon the exterior. As to the external entity, the network element hides the internal structure, equivalent to a router or switch. In Fig. 1, Fi/f is external interface of FE, Fi is interface between CE manager and FE manager. Ff is interface between FE manager and FE, Fc is interface between CE manager and CE, Fr is interface between CE and CE. Fi is interface between FE and FE, Fp is interface between CE and FE, the communication standard of Fp is the core research content of ForCES protocol (Zhuge *et al.*, 2011).

In ForCES router, the transmission of protocol messages between CE and FE is the key in restricting the communication performance of the whole communication system. Therefore, the choice of transport protocols is related to the performance of the whole system. In RFC 5811, SCTP has been defined as the transport protocol of TML. SCTP has most of the required characteristics to construct a robust TML, support a simpler coding and reduce many problems such as interoperability. So, the next research and implementation of this study is based on SCTP network.

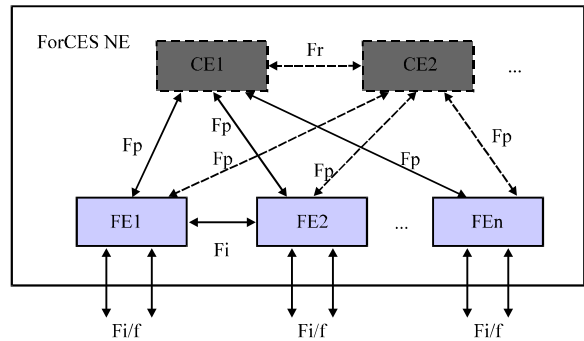


Fig. 1: Architecture of ForCES Network

ForCES architecture included two parts, Protocol Layer (PL) and TML (Doria *et al.*, 2010). PL is responsible for generating ForCES protocol messages independent of the type of TML. At the same time, PL manages and controls TML. TML is responsible for the transmission of PL messages and TML shields transmission medium component when messages are transmitted. It provides a transparent and safe message transmission channel for PL. TML underlies under PL, TML includes CE TML of control plane and FE TML of data plane. The communication between CE PL and FE PL is realized by CE TML and FE TML. When there are messages to be transmitted, PL will send ForCES messages to TML of itself. Then to realize the transmission of ForCES messages between two PL entities through the message exchange between two TMLs.

The CE TML router will communicate with multiple or even hundreds of FE TMLs at the same time in ForCES framework router. This is a typical one-to-many communication mode. Therefore, the implementation of TML communication mainly adopts server/client mode, namely take CE TML as a server, FE TMLs as the clients. FE TML issues a task request to CE TML and they establish a connection and take communication processing. In this mode, CE TML chooses the communication mechanism to realize and coordinate the communication flow and it is the key research (Qaddoura *et al.*, 2002).

**BANDWIDTH ALLOCATION IN FORCES TML**

In ForCES router, the TML is responsible for the transmission of ForCES protocol messages in SCTP transmission medium. In order to transmit various protocol messages efficiently, TML need take proper bandwidth allocation for each sub channel in ForCES channels. As what showed in Fig. 2, based on TCP/IP network, ForCES TML provides services for protocol layer. ForCES TML is also responsible for the

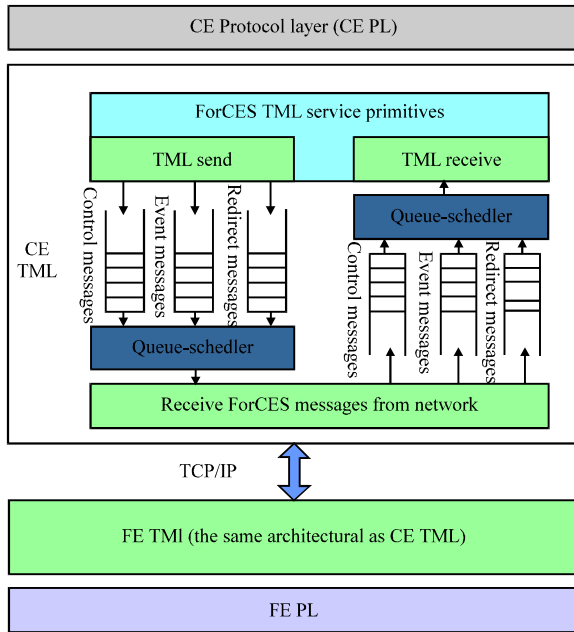


Fig. 2: Architecture of TML

transmission of ForCES protocol messages between protocol layers. ForCES protocol layer will produce three types of ForCES protocol messages, control protocol messages, event protocol messages and redirection protocol messages. And, control protocol messages are generated by FEs or CEs for control information exchange. Such as the protocol messages for configuring or querying the parameters of FEs or CEs. The normal transmission of control protocol messages is the necessary condition for the normal work of the whole router. Event protocol messages are generated when the event registered in the event dispatch module happens. Such as, control channel messages arrival event, event channel messages arrival event, redirect channel messages arrival event, listening socket connection events. And messages will be generated when these events occurs. These types of event messages can be cumulated and CE can get the event message through sending a query to FE. So, this event message can tolerate loss. Redirect protocol messages are used to load the data packets. For example, FEs need redirect network packets which come from the network via the network interface, to CEs. CEs also need send a number of data packets which are generated by these CEs, to FEs. Then FEs send these packets to the exterior of the network. And the data is transmitted by the redirect protocol messages of ForCES protocol. Therefore, the ForCES TML must ensure the transmission of control protocol messages. To transmit event protocol messages and redirect protocol

messages effectively without affecting the transmission of control protocol messages (Haleplidis *et al.*, 2010; Stoica *et al.*, 2003; Sharma *et al.*, 2007).

The focus of traditional bandwidth allocation algorithm is to improve distribution efficiency, and pay little attention to message attributes and the importance of different messages. So, it lacks pertinence and encounters a bottleneck in improving utilization rate of cybersource. According to different characteristics of different message flows in ForCES router, this study puts forward a new bandwidth allocation mechanism. It is rte and buffer based active queue management-dynamic probabilistic priority scheduling in ForCES channel.

**Summarize of RBAQM-DPP algorithm:** RBAQM-DPP algorithm model is shown in Fig. 3. In this algorithm, firstly, set up a FIFO queue, respectively for the control messages, the event messages and the redirect messages. At the same time, set the initial scheduling probability for the control message queue, the event message queue and the redirect message queue. Note the probabilities as  $p_c$ ,  $p_e$ ,  $p_d$ , the sum of the three probabilities is 1. Then, set  $Min_{thd}$  as the minimum buffer threshold of queue and set  $Max_{thd}$  as the maximum buffer threshold of queue for each message queue, respectively. Affect the allocation of link bandwidth resource by adjusting the queue scheduling probability.

In the module of priority scheduling in probability, according to values of  $p_c$ ,  $p_e$ ,  $p_d$  to schedule forwarding packets from the head of each message queue. Obviously, in the event that the total link bandwidth is fixed, if  $p_c$  increases, the corresponding bandwidth of control channel will increase. It will lead to diminished occupied bandwidth of the event channel and redirect channel. In RBAQM-DPP, the value variability of  $Min_{thd}$  and  $Max_{thd}$  provides more initiative and plasticity for the algorithm. In different stages of the transmission, administrators can better meet the transmission requirement of all types of messages by setting different values of  $Min_{thd}$  and  $Max_{thd}$ . So, as to further improve the quality of transmission (Guo, 2001).

**Realization of RBAQM-DPP algorithm:** The algorithm consists of two techniques, active queue management and queue scheduling. Queue scheduling mainly realizes the rational distribution on link bandwidth resource through the scheduling mechanism, so as to ensure the reserved bandwidth strictly. Active queue management adopts some strategy to decide whether to discard the message packet according to feedback information of queue in the arrival of packet (Athuraliya *et al.*, 2001; Jaseem *et al.*, 2010; Kunniyur and Srikant, 2000).

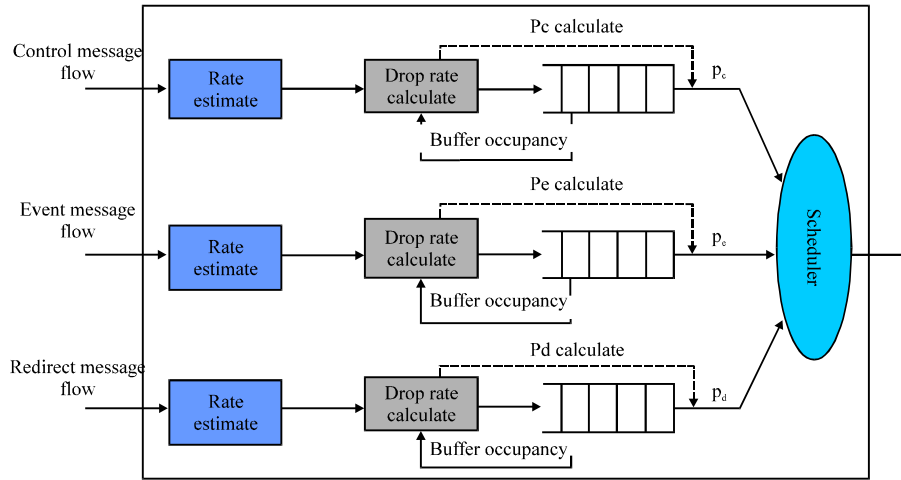


Fig. 3: RBAQM-DPP Scheduling model

**Queue scheduling:** When the algorithm executes, firstly, according to the initial values of scheduling probability, to schedule and forward the message packets of each channel. In the service of each round, the probability of scheduling and forwarding control messages is  $p_c$ . The probability of scheduling and forwarding event messages is  $p_e$ . And the probability of scheduling and forwarding redirect messages is  $p_d$ . Thus, the occupied bandwidth of three message channels is:

$$B_c = p_c \times B, B_e = p_e \times B, B_d = p_d \times B \quad (1)$$

where,  $B$  is total link bandwidth between CE and FE communication.  $B_c, B_e, B_d$  are, respectively the channel bandwidth of the control channel, the event channel and the redirect channel.

**Active queue management:** In the transmission process of news groups, usually to improve the utilization of network bandwidth by taking queue cache (Matsurfum and Aibara, 1999). Active queue management is important means to realize the control of network congestion. It influences the link bandwidth allocation through the control on buffer resource. On a single network node, its control goal is to improve the throughput and allocate bandwidth resources reasonably at the same time (Bennett and Zhang, 1996; Hollot *et al.*, 2001).

Firstly, to assign the limited bandwidth to three different message flows effectively through reasonably setting  $Min_{thd}$  and  $Max_{thd}$  of each message queue. Therefore, when the message starts to lose packets with a probability, the determination of  $Min_{thd}$  and  $Max_{thd}$  of this queue determines the lower and upper occupied bandwidth. To set values of  $Min_{thd}$  and  $Max_{thd}$  of the

queue by the characteristics observed from the three actual ForCES message flows and priority of the three kinds of messages (Bennett and Zhang, 1997; Floyd and Jacobson, 1993).

Then, when there is message going into the queue, first of all, obtain the buffer occupancy information of this kind of message queue. Compare the obtained results with the prior given threshold of queue length (Ott *et al.*, 1999). When the message queue length exceeds  $Max_{thd}$  of the queue, it shows that the message channel has congestion phenomenon, then discard the message directly. When the message queue length is less than  $Max_{thd}$ , the enqueueing module will firstly calculate the packet loss probability  $p_{q-dp}$  of current message. The loss probability  $p_{q-dp}$  is caused by the changing queue length. Then, according to the average arrival rate of message flows in each channel, calculate packet loss coefficient  $p_{r-dp}$  to determine whether to perform loss. Packet loss coefficient  $p_{r-dp}$  is caused by the changing average arrival rate of the three kinds of message flows. That means that the total loss rate is:

$$P_{dp} = P_{q-dp} \cdot P_{r-dp} \quad (2)$$

Calculate  $p_{q-dp}$  and  $p_{r-dp}$  by the following method:

Through the above analysis, we know that when the message queue length exceeds  $Max_{thd}$ , discard the message directly, so,  $p_{q-dp} = 1$ . When the message queue length is less than the minimum buffer threshold  $Min_{thd}$ , it shows that the news channel has no congestion. And all the arrival messages are permitted to enter the current message queue, so  $p_{q-dp} = 0$ . When the message queue length is between  $Max_{thd}$  and  $Min_{thd}$ , calculate the packet loss rate of a new arrival message packets like Eq. 3:

$$p_{q\_dp} = a(1-t)^2 \tag{3}$$

In Eq. 3,  $a$  is a constant,  $l$  is the current message queue length,  $\text{Min}_{thd} < l < \text{Max}_{thd}$ ,  $t = \text{Min}_{thd}$ . In current study, there are also a lot of methods using the linear increase, and at  $\text{Max}_{thd}$  it reaches the maximum loss value. It can be said that this issue is still a relatively open problem. Solve the derivation on the left and right of Eq. 3:

$$p_{q\_dp} = 2al - 2at \tag{4}$$

In Eq. 4, the value of  $a$  determines the smoothing degree of the curve. In addition, with the increasing of queue length  $l$ , the increasing rate of packet loss rate increases. We can try to maintain the queue in a stable queue which receives burst message flows. Now, a lot of studies propose that the transition from discarding to discarding completely should tend to be smooth as soon as possible. Namely, when the queue length reaches  $\text{Max}_{thd}$ , the value of  $p_{q\_dp}$  is close to 1. Thus, once we determine the values of  $\text{Max}_{thd}$  and  $\text{Min}_{thd}$ , the value of  $a$  is determined.

In short:

$$p_{q\_dp} = \begin{cases} 0, & l \leq \text{Min}_{thd} \\ a(1-t)^2, & \text{Min}_{thd} < l < \text{Max}_{thd} \\ 1, & l > \text{Max}_{thd} \end{cases} \tag{5}$$

According to average arrival rate  $r$  of message stream, the enqueue module determines packet loss coefficients  $p_{r\_dp}$ . Using  $t_k$  and  $l_k$  to express arrival time and length of the  $k$ th message packet, respectively. When a new message arrives, estimate the average arrival rate of the message stream using exponential average algorithm:

$$r_{new} = (1-\alpha) \frac{l_{k+1}}{T_{k+1}} + \alpha r_{old} \tag{6}$$

There,  $\alpha = e^{-T_{k+1}/K}$  is index weight,  $r_{old}$  is the average arrival rate of message flow before the arrival of the new message packet.  $T_{k+1} = t_{k+1} - t_k$ , it shows the arrival interval between the  $(k+1)$ th message and the  $k$ th message,  $K$  is a constant.

The index weight  $\alpha$  is a variable which is different from the index weight of constant, arrival interval time  $T$  of packets is a variable. The intent of using such a variable index weight is to reflect the average process of the message flow independent of the packet structure more accurately. More precisely, if the index weight is a constant, the estimated average rate of the arrival message flow is sensitive to the message packets length.

And the estimated rate is different from real-time arrival rate because of some reason. Contrarily, using such an index weight of variable  $\alpha$ , the estimated rate will asymptotically tend to be the real rate.

In the expression  $\alpha = e^{-T_{k+1}/K}$  of  $\alpha$ , give many tradeoffs to the value of  $K$ . Firstly, a smaller  $K$  can increase the response and produce rapid rate fluctuations, while a larger  $K$  can better avoid potential instability of the system. Secondly, delay jitter changes the arrival time interval of message packets. This will lead that the difference between the estimated rate and the real-time rate increases ceaselessly. In order to overcome this effect, according to the experience, the value of  $K$  and the delay of message flow should be in the same order of magnitude. In addition, the value of  $K$  can't exceed the average duration of message flow. In view of the above constraints, a suitable value of  $K$  should be set between 100 and 500 msec. Known from the above, the average arrival rate of the current message flow is  $r$ . Note the average arrival rate of the current control messages as  $r_c$ , the average arrival rate of the event messages as  $r_e$ , the average arrival rate of the redirect messages as  $r_d$ . So the loss coefficient is:

$$p_{r\_dp} = \begin{cases} 0, & r_c + r_e + r_d < B \\ 1, & r_c + r_e + r_d \geq B \end{cases} \tag{7}$$

Put Eq. 7 into Eq. 2 to get:

$$p_{dp} = \begin{cases} 0, & r_c + r_e + r_d < B \\ p_{q\_dp}, & r_c + r_e + r_d \geq B \end{cases} \tag{8}$$

Next, as to all the reaching news, we will adjust the value of scheduling probability of the kind of message according to the value of  $p_{dp}$  and the message queue length. As to the control channel, note that the current queue length is  $l_c$ , the total loss rate is  $p_{dp}^c$ . When the control message packet arrives, the packet loss rate of control message packet caused by the queue length is  $p_{q\_dp}^c$ . Calculate the scheduling probability  $p_c$  of control channel according to  $l_c$  and  $p_{dp}^c$ , get from the Eq. 5:

$$p_{dp}^c = \begin{cases} 0, & r_c + r_e + r_d < B \\ p_{q\_dp}^c, & r_c + r_e + r_d \geq B \end{cases} \tag{9}$$

$$l_c = \begin{cases} 0, & p_{dp}^c = 0 \\ \text{Max}_{thd}, & p_{dp}^c = 1 \\ l_c, & 0 < p_{dp}^c < 1 \end{cases} \tag{10}$$

So, it can be  $p_c$ :

$$p_c = \frac{r_e \times (1 - p_{dp}^e) + \beta l_e}{B} \quad (11)$$

And,  $\beta$  is the corresponding proportion coefficient caused by the queue length and the bandwidth,  $\beta$  is a fixed value.

As to the event channel, note that the current queue length is  $l_e$ , the total loss rate is  $p_{dp}^e$ . When the event message packet arrives, the packet loss rate of event message packet caused by the queue length is  $p_{q-dp}^e$ . Adjust the scheduling probability according to  $p_c$  and  $p_{dp}^e$ , adjustment process is as follows:

The occupied bandwidth of control channel is  $B_c$ , so the residual bandwidth is  $B - B_c$ , get from Eq. 5 similarly:

$$p_{dp}^e = \begin{cases} 0, & r_e + r_d < B - B_c \\ p_{q-dp}^e, & r_e + r_d \geq B - B_c \end{cases} \quad (12)$$

$$l_e = \begin{cases} 0, & p_{dp}^e = 0 \\ \text{Max}_{\text{th}}, & p_{dp}^e = 1 \\ l_e, & 0 < p_{dp}^e < 1 \end{cases} \quad (13)$$

Therefore, remove the bandwidth of control channel, the relative scheduling probability  $p_e$  and  $p'_d$  can be gotten from the remaining bandwidth  $B - B_c$ . And  $p'_e$  is relative scheduling probability of the event channel,  $p'_d$  is relative scheduling probability of the redirect channel. They are:

$$p'_e = \frac{r_e \times (1 - p_{dp}^e) + \beta l_e}{B - B_c} \quad (14)$$

$$p'_d = \min \left( (1 - p'_e), \frac{r_d}{B - B_c} \right) \quad (15)$$

In the remaining bandwidth  $B - B_c$ , note that the relative surplus scheduling probability is  $\Delta p'_c$  in addition to the event channel and the redirect channel:

$$\Delta p'_c = \max(0, 1 - p'_e - p'_d) \quad (16)$$

Thus, it obtains:

$$p_e = p'_e \times (1 - p_c) \quad (17)$$

$$p_d = p'_d \times (1 - p_c) \quad (18)$$

To except the event channel and the redirect channel, the remaining scheduling probability  $\Delta p_c$  by calculating is:

$$\Delta p_c = \Delta p'_c \times (1 - p_c) \quad (19)$$

To obtain the final,  $p_c$  is:

$$p_c = p_c^{\text{old}} + \Delta p_c \quad (20)$$

In the bandwidth  $B - B_c$ , if the sum of occupied bandwidth of the event channel and the redirect channel is less than  $B - B_c$ , allocate the residual bandwidth to control channel. And this will provide better protection for the transmission of control message.

Therefore, because of the RBAQM-DPP scheduling mechanism, TML can improve the utilization of network bandwidth. At the same time, it guarantees the transmission of control messages and avoids the starved phenomenon of the low priority channel. So it improves the communication performance of ForCES router fundamentally.

### RESEARCH ON FLOW CONTROL MECHANISM BETWEEN FORCES CHANNELS

The research in the above section is in the case of a CE connected with a single FE. And it is the bandwidth allocation algorithm that among the three news channels of the ForCES channels. FE can occupy the full bandwidth of CE communication link naturally. But usually, in a ForCES router, one CE generally controls dozens or even hundreds of FEs, as shown in Fig. 4. When a CE connects with multiple FEs, the total bandwidth  $B$  allowed by CE communication will be assigned to more than one FE. That's because it wants to take normal data interaction between the FEs and CE, especially to ensure the transmission of high priority messages in ForCES channels. When the sum of message rate in each ForCES channel is larger than the total bandwidth of CE communication, if we don't take any measures, it will inevitably cause congestion. It results in the message packets of each ForCES channel discarding randomly at the connection nodes of CE and FEs. In the case of severe congestion, it might even lead that the messages of a ForCES channel were discarded totally or the control messages were largely abandoned. This would ceaselessly increase the delay time of message transmission (Kanher and Sethu, 2001).

Obviously, good control of the ForCES channel flow is very important for the accurate and safe message interaction between CE and multiple FEs. This requires a more accurate prediction for the flow of each ForCES channel. And the most important is how to forecast the average arrival rate of ForCES channel messages in the next time. The forecasting precision will directly affect the accuracy of congestion judgment and whether the link bandwidth resources can be fully utilized.

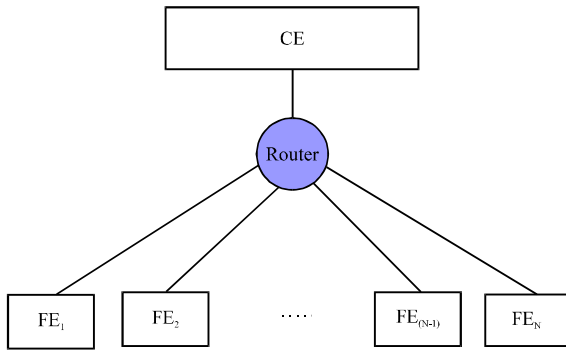


Fig. 4: Simple ForCES router model

The ForCES router has good expansibility. According to the extension mechanism of ForCES router, add flow control module in the CE end to achieve flow control mechanism. It is responsible for the over-all planning of flow control mechanism. Add a component in the FE end, abstract the flow control function resource, finish the query configuration channel which is provided for the top layer and achieve the function entity of flow control. The model of flow control mechanism is as shown in Fig. 5. Therefore, according to each ForCES channel, to design an algorithm called Prediction-based Flow Control and Configuration (PFCC) in the CE terminal.

**Prediction-based flow control and configuration algorithm in ForCES channels:** The core idea of PFCC algorithm is that, firstly, predict accurately the message flow of each ForCES channel in the next time. Then CE allocates available bandwidth for each ForCES channel instantaneously, to ensure message transmission quality and realize the reasonable allocation of cybersource at the same time.

**Flow prediction of ForCES channels:** In the ForCES router, there are a large number of observations about the changes of message flows in each channel. It can be found that, the message flows of each ForCES channel don't show a strong long range correlation. Thus, using Minimum Mean Square Error (MMSE) to predict is more precise. And MMSE algorithm is simple, has a small amount of calculation which meets the demand of on-line prediction. Research shows that, MMSE algorithm has more accurate prediction for aggregation flows than the single flow. And the flow prediction for a ForCES channel includes three converged sub message flows, so the prediction effect is more ideal.

Next, describe how to use MMSE algorithm to predict the arrival rate of the next packet.

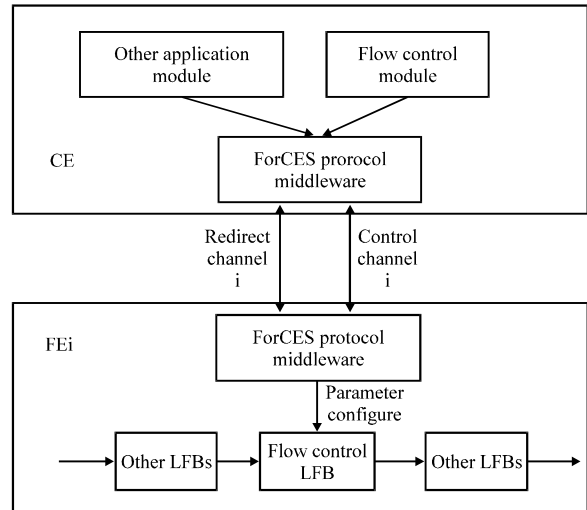


Fig. 5: Model of flow control mechanism

Using  $\{X_t\}$  represent linear stochastic process of the sent packet rate in  $t$  time, the arriving packet rate in the next time represents by the linear combination of the sent packets:

$$X_{t+1} = \mu_1 X_t + \mu_2 X_{t-1} + \dots + \mu_p X_{t-p+1} + \varepsilon = \mu X + \varepsilon \quad (21)$$

There into,  $\mu = (\mu_1, \mu_2, \dots, \mu_p)$  is partial correlation coefficient,  $X = (X_t, X_{t-1}, \dots, X_{t-p+1})$ ,  $\varepsilon$  is estimation error,  $p$  is order number.

Note the estimated value of  $X_{t+1}$  as  $\hat{x}_{t+1}$ , the estimated value of  $\mu$  as  $\hat{\mu}$ .

In order to make the criterion error of estimation rate the minimum, according to the MMSE algorithm that  $E[\varepsilon^2]$  takes the minimum value, namely solve the minimum value of:

$$E[\varepsilon^2] = E[(X_{t+1} - \hat{x}_{t+1})^2]$$

Make:

$$\frac{\partial E\varepsilon^2}{\partial \mu_k} = 0, k=1, \dots, p$$

thus, it obtains:

$$\hat{\mu} = [r_1, r_2, \dots, r_p] \begin{bmatrix} 1 & r_1 & r_2 & \dots & r_{p-2} & r_{p-1} \\ r_1 & 1 & r_1 & \dots & r_{p-3} & r_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p-2} & r_{p-3} & r_{p-4} & \dots & 1 & 1 \\ r_{p-1} & r_{p-2} & r_{p-3} & \dots & r_1 & 1 \end{bmatrix}^{-1} \quad (22)$$



And,

$$r_t = \frac{1}{p} \sum_{i=t+1}^p X_i X_{t-i}$$

Thus, the rate of the arriving packet message in the next time can be written as:

$$X_{t+1}^{\wedge} = [X_t, X_{t-1}, \dots, X_{t-p+1}] \mu^T \quad (23)$$

Thus, for a single ForCES channel, using the index average algorithm to calculate the average rate of the channel message when the next packet arrives:

$$r_k^{new} = (1 - \sigma) X_{t+1}^{\wedge} + \sigma r_k^{old} \quad (24)$$

And,  $r_k^{old}$  is the average transmission rate of messages in the current ForCES channel. And  $r_k^{new}$  is the average transmission rate of messages in the ForCES channel when the number  $t+1$  packet arrived.

**Overview of PFCC algorithm:** When the algorithm executes, firstly set the initial bandwidth  $B_k^{ini}$  for the corresponding ForCES channel of each FE which is connected with CE. The allocation bandwidth is  $B_k^{con}$ , the minimum bandwidth is  $B_{min}$  and the maximum bandwidth is  $B_{max}$ . There into,  $B_k^{ini}$  is the initial allocation bandwidth for the  $k$ th FE (ForCES channel),  $B_{kcon}$  is the allocation bandwidth for the  $k$ th ForCES channel. CE gets  $N$  which is the number of FEs connected with it every certain period and the current message transmission rate  $r_k$  of each ForCES channel. Because there will be new FE to establish a connection with CE or the established connection disconnects, so  $N$  is real-time changing. In addition, actual occupied bandwidth values of ForCES channel may real-time vary with the arrival rate of the message flow. The algorithm adjusts allocation bandwidth  $B_k^{con}$  of each ForCES channel according to the change of  $N$  and the current actual occupied bandwidth values  $B_k^{rel}$  of ForCES channel.

However, in PFCC, through reasonable adjustment of allocated bandwidth  $B_k^{con}$  for ForCES channel, the total allocated bandwidth of channels should be within the allowed bandwidth of CE communication. Thus it can avoid congestion. For a single ForCES channel, once the allocated bandwidth  $B_k^{con}$  is determined. The premise is that the total bandwidth of all the sub-channels of the single channel is within  $B_k^{con}$ . We can achieve this in accordance with RBAQM-DPP algorithm and make sure that the end-to-end delay of various messages is under control.

**The realization of PFCC algorithm:** Firstly, the same initial bandwidth  $B_{kmi}$  should be allocated for the ForCES channel of each FE which is based on the quantity  $N$  of FE connected to CE. The method to calculate  $B_k^{ini}$  is as follows:

$$B_k^{ini} = \frac{B}{N \times p_t} \quad (25)$$

And  $p_t$  is the probability that  $N$  FEs have message interaction with CE at the same time. At any time, the probability that each FE has message transmission with CE is  $\gamma$  and each FE is independent with others. So, the probability that  $N$  FEs have message transmission with CE at the same time is  $p_t = \gamma^N$ . That is:

$$B_k^{ini} = \frac{B}{N \times \gamma^N} \quad (26)$$

So, the initial bandwidth of each FE can be expressed as:

$$\begin{bmatrix} FE_1 \\ FE_2 \\ \vdots \\ FE_N \end{bmatrix} = \begin{bmatrix} B \\ N \times \gamma^N \end{bmatrix} \quad (27)$$

Suppose that  $B_{max} = B_k^{ini} \cdot \theta$ ,  $B_{min} = b$ , where  $0 < \theta < 1$  and  $b > 0$  which is a constant. We can adjust the values of  $\theta$  and  $b$  according to the actual situation. The allocated bandwidth  $B_k^{con}$  is initialized as  $B_{max}$ .

Then, the value of  $B_k^{con}$  is adjusted according to the current bandwidth  $B_k$  of ForCES channel and the average arrival rate of messages in ForCES channel in the next moment. And the value of  $B^{rel}$  is:

$$B^{rel} = B_1^{rel} + B_2^{rel} + \dots + B_k^{rel} + \dots + B_N^{rel} = \sum_{k=1}^N B_k^{rel} \quad (28)$$

Then, the total spare bandwidth  $\Delta B$  at present is:

$$\Delta B = B - B^{rel} = B - \sum_{k=1}^N B_k^{rel} \quad (29)$$

Next, we can get the average rate of messages in ForCES channel when the next package arrives to re-determine the allocated bandwidth.

In the last section, when the  $(t+1)$ th package arrives, the average arrival rate is:

$$r_k^{t+1} = (1 - \sigma) X_{t+1}^{\wedge} + \sigma r_k^t \quad (30)$$

Thus, when the  $(t+1)$ th package is sent, the sum of average rate of messages in each ForCES channel is:

$$r^{t+1} = r_1^{t+1} + r_2^{t+1} + \dots + r_k^{t+1} + \dots + r_N^{t+1} = \sum_{k=1}^N r_k^{t+1} \quad (31)$$

If  $r_k^{t+1} = 0$ , then  $B_k^{\text{con}} = B_{\min}$ . The reason why the allocated bandwidth isn't set to zero when there is no message to transmit is because of the burst stream. So, we can transmit data before the bandwidth is allocated and improving the transport efficiency of burst data stream.

If  $r_k^{t+1} > 0$ :

$$B_k^{\text{con}} = B_k^{\text{nl}} + \Delta B \times \frac{r_k^{t+1}}{r^{t+1}} \quad (32)$$

Now, the allocated bandwidth for each FE is:

$$\begin{bmatrix} \text{FE}_1 \\ \text{FE}_2 \\ \vdots \\ \text{FE}_N \end{bmatrix} = \begin{bmatrix} B_1^{\text{nl}} + \Delta B \times \frac{r_1^{t+1}}{r^{t+1}} \\ B_2^{\text{nl}} + \Delta B \times \frac{r_2^{t+1}}{r^{t+1}} \\ \vdots \\ B_N^{\text{nl}} + \Delta B \times \frac{r_N^{t+1}}{r^{t+1}} \end{bmatrix} \quad (33)$$

That is, the total spare bandwidth is allocated to each FE according to the proportion of message transmission rates.

If  $r_k^{\text{con}} > B_{\text{max}}$ , then set  $B_{\text{con}} = B_{\text{max}}$ , so that,  $B_k^{\text{con}}$  will not be so large. The aim is that even if the channel is subjected from malicious attacks, for the entire network bandwidth  $B$ , the bandwidth expense is within  $B_{\text{max}}$ . So it won't cause congestion of the entire network. In conclusion, in PFCC, a reasonable bandwidth allocation for each ForCES channel can help users to avoid congestion. And it also can offer a reliable guarantee for end-to-end delay.

## THE SIMULATION AND TEST

**Simulation and test of RBAQM-DPP algorithm:** Use NS2 network simulator to simulate the work of FE and CE in ForCES framework router under SCTP networks. To choose different news source models and strict probability priority scheduling algorithm (SP) recommended by RFC5812. Then compare and verify the performance of RBAQM-DPP algorithm proposed by this study.

The network simulation topology which is built in test is as Fig. 6, CE (node 0) and FE (node 2) are connected via a router (node 1). Three sources (node 3, 4, 5) are, respectively used to simulate the actual

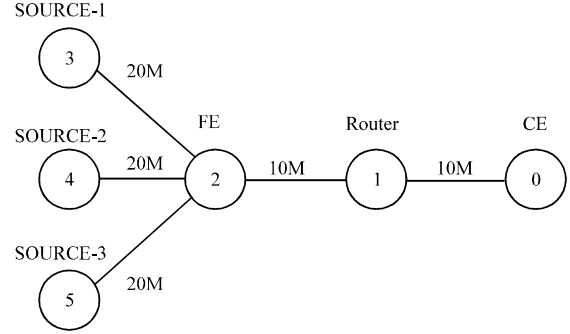


Fig. 6: Topology of network simulation

control messages, event messages and redirect messages. Message queue management and scheduling between FE and the router respectively use RBAQM-DPP and SP algorithm. In the CE terminal, to receive and analyze the transmission efficiency of each kind of information forwarded from FE.

In the test, the link bandwidth is as what showed in Fig. 6, set the link delay to 0.1 msec, the queue length to 100,  $K = 300$  msec.

- Each message source sends messages at a constant rate, set the control messages in high priority channel (HP) at the rate of  $4 \text{ Mb sec}^{-1}$ . Set the event messages in middle priority channel (MP) at the rate of  $3 \text{ Mb sec}^{-1}$ , the redirect messages in low priority channel (LP) at the rate of  $2 \text{ Mb sec}^{-1}$ , set the sampling time to 0.05 sec

By the two Fig. 7 and 8 when each source rate is less than the link bandwidth 10 Mb. RBAQM-DPP, SP, the control messages, the event messages and the redirect messages all have normal transmission. RBAQM-DPP and SP algorithm essentially have the same effect.

- Set that the source model is that rate increases with time, the sampling time is 0.05 sec. To observe the performance of RBAQM-DPP and SP algorithm when there is congestion

From the Fig. 9 and 10 it can be seen that, when the news source rate is small, three kinds of message flows in RBAQM-DPP and SP don't have packet loss. However, as the news source rate increasing and congestion happening, the event messages and the redirect messages in SP are in severe packet loss phenomenon. Redirect channel is even in starved phenomenon. While in the action of RBAQM-DPP, even congestion happens, the event messages and the redirect messages can still maintain a certain transmission rate. It won't be in starved phenomenon in low priority channel.

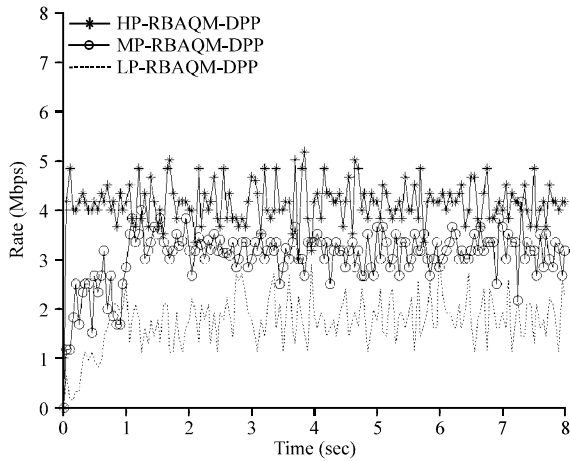


Fig. 7: Performance of RBAQM-DPP without congestion

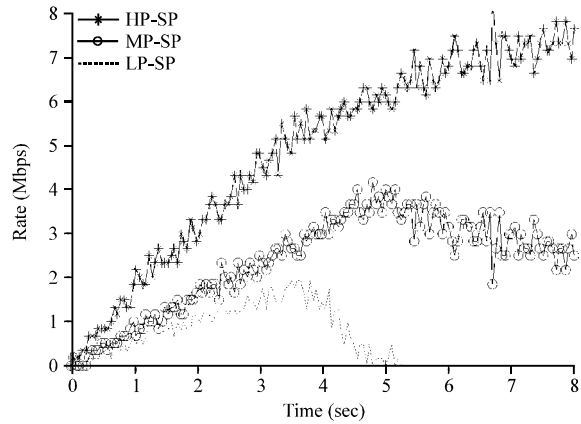


Fig. 10: Performance of SP in congestion

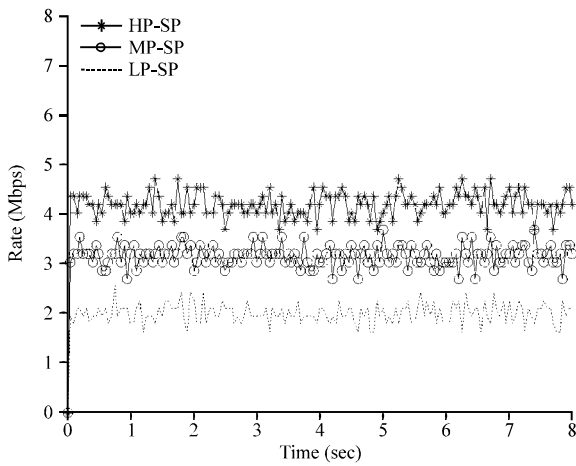


Fig. 8: Performance of SP without congestion

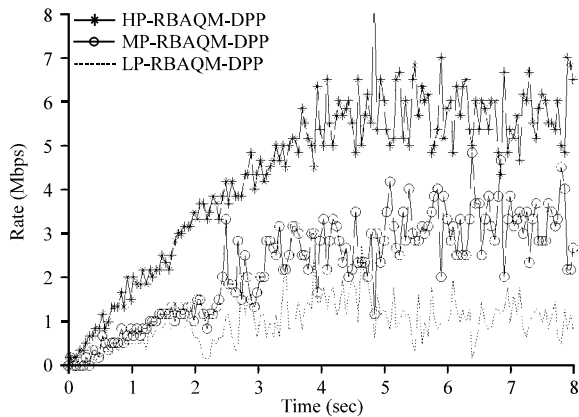


Fig. 9: Performance of RBAQM-DPP in Congestion

- In RBAQM-DPP, through taking different values of  $Max_{thd}$  and  $Min_{thd}$ , to observe the output bandwidth of each message stream, sampling time is 0.5 sec. In order to better reflect the test results, set the output of each message source at a constant rate. Set the rate of the control messages at  $8 \text{ Mb sec}^{-1}$ , the rate of the event messages at  $3 \text{ Mb sec}^{-1}$ , the rate of the redirect messages at  $4 \text{ Mb sec}^{-1}$

Such as shown in the Fig. 11 and 12, different values of  $Max_{thd}$  and  $Min_{thd}$ , the output bandwidth ratio of each message channel has a more obvious change. To increase the values of  $Max_{thd}$  and  $Min_{thd}$ , the output bandwidth of control channel is increased. When  $Min_{thd} = 30$ ,  $Max_{thd} = 75$ , the output bandwidth of the control channel nearly reaches the input rate  $8 \text{ Mb sec}^{-1}$  of the control message. To reduce the values of  $Max_{thd}$  and  $Min_{thd}$ , the output bandwidth of the control channel is decreased. But the output bandwidth of the event channel and the redirect channel increases. This characteristic of RBAQM-DPP provides more workable space for network administrators. According to the change of different message flows number in each stage, the administrator can dynamically adjust the value of the queue length threshold. For example, at the beginning of establishing a connection, the control messages in message stream apparently accounts for the major part. So, the maximum threshold and the minimum threshold should be relatively set to large values. But once connects, the control messages proportion in message flow accounts for only a small part, the main part is redirect messages data. So, the maximum and minimum threshold should be relatively set to small values, so as to improve the communication performance of the whole system.

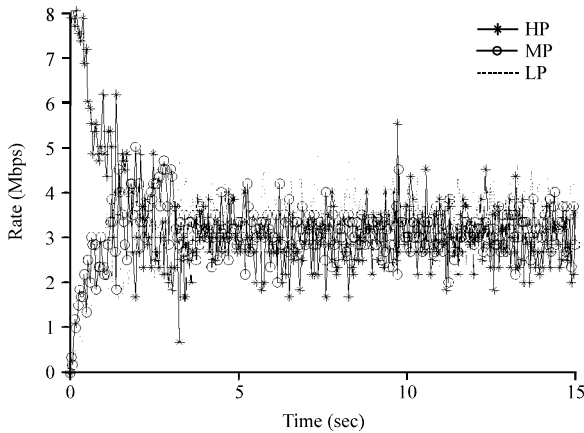


Fig. 11:  $\text{Min}_{\text{thd}} = 20, \text{Max}_{\text{thd}} = 50$

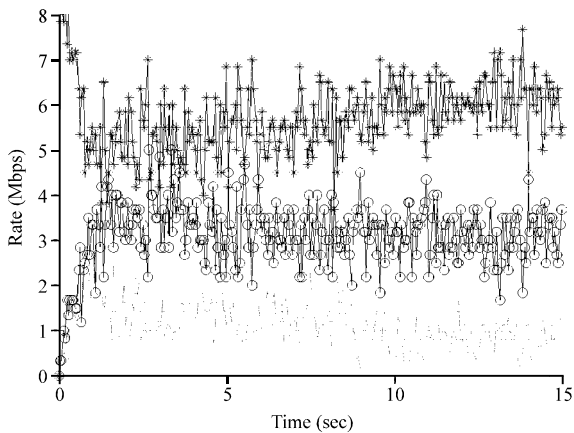


Fig. 12:  $\text{Min}_{\text{thd}} = 30, \text{Max}_{\text{thd}} = 75$

**Test and Analysis of Multi-FE Flow control algorithm:**

In the NS2 network simulation tool, to set up network simulation environment as shown in Fig. 13, messages of FE-1 (node 2), FE-2 (node 3) and FE-3 (node 4) converge to CE (node 0) through Router (node 1). The node 6, 7, and 8 are message sources, each source can agent various types of news. The source agents generate control messages, event messages and redirect messages.

- The link bandwidth settings were as shown in Fig.13, the sending rate of each message source increases with time. Observe the transmission of each message in the CE end when using PFCC algorithm.

As shown in the Fig. 14, when the sum of the three news source rates is less than  $10 \text{ Mb sec}^{-1}$ , the output bandwidth of each FE message increases when the message source rate increases. With the source rate increases ceaselessly, around 7 sec, the sum of each source rate has been larger than  $10 \text{ Mb sec}^{-1}$ . But under

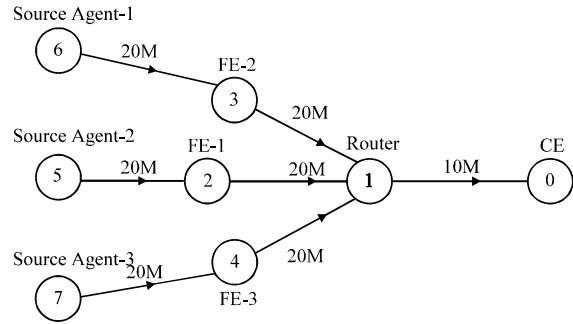


Fig. 13: Network simulation environment

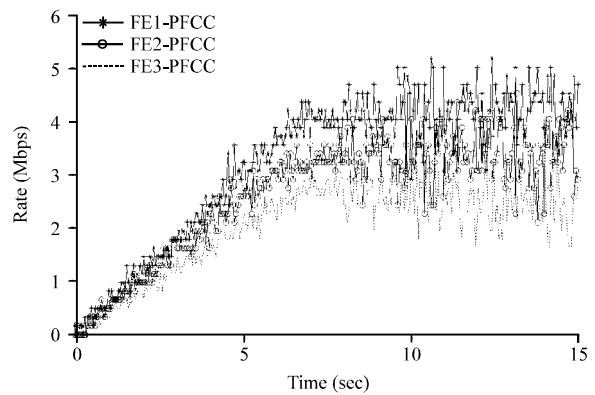


Fig. 14: Output bandwidth of each FE Message in PFCC

the effect of PFCC algorithm, it avoids congestion on node 1. Each FE message is maintained at the output bandwidth of the critical congestion, there won't be messages abandoned largely phenomenon in a FE.

- Take FE-1 for example, to set the messages source to output with constant speed, draw a comparison on the output bandwidth of each message stream when with PFCC algorithm and without PFCC algorithm in FE-1.

Seen from the Fig. 15 and 16 when there is no PFCC algorithm, the transmission situation of the three kinds of ForCES protocol news in FE-1 is very chaotic. For much time the control messages occupies almost no output bandwidth, which undoubtedly increases the delay time of control messages. And when there is PFCC algorithm, in the available bandwidth of FE-1, various news especially the control messages can maintain stable output bandwidth. It won't exhibit abandoned largely phenomenon. Thus the QoS of various messages can get reliable guarantee, to provide a good network transmission environment for users.

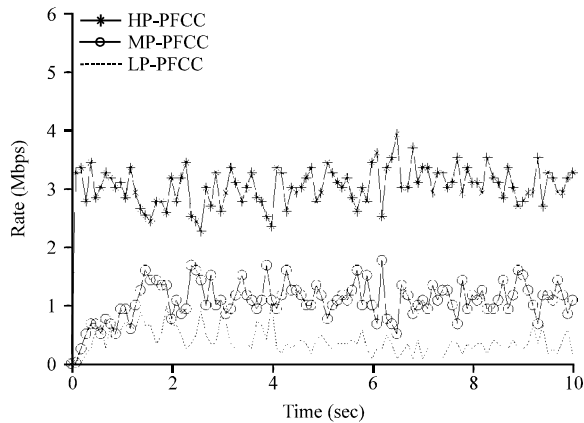


Fig. 15: Output bandwidth of each message stream with PFCC in FE-1

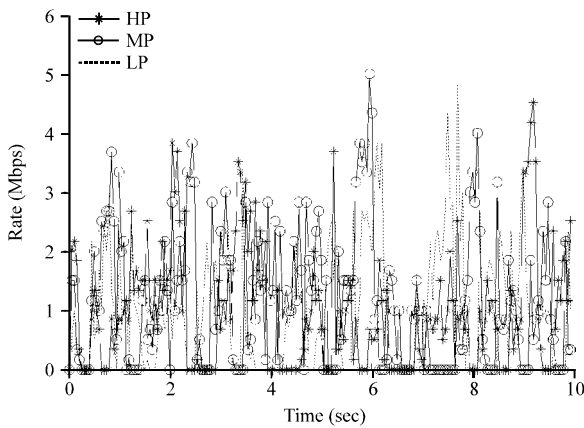


Fig. 16: Output bandwidth of each message stream without PFCC in FE-1

**CONCLUSION**

This study studies the characteristics of bandwidth allocation in ForCES transport mapping layer and puts forward RBAQM-DPP algorithm. The algorithm calculates the rejection probability, according to the message arrival rate and queue length when messages enqueue. Then adjust the dequeue probabilities according to the rejection probability, dynamically change the bandwidth of message queue. RBAQM-DPP algorithm can well protect the control message flow for single FE and ensure the transmission of the control message flows. It improves the transmission rate of the event message flows and the redirect message flows. However, under normal circumstances, one CE is often associated with multiple FEs. There is also congestion on the connection nodes between CE and FE which is caused by uneven or too excessive traffic load among the FEs. Thus, this study

proposes PFCC algorithm among the ForCES channels. In PFCC flow control algorithm, CE can reasonably allocate the total bandwidth to each FE according to real-time traffic conditions of FEs. So it realizes the flow control in the internal channels of the ForCES router. And it guarantees the transmission efficiency of data flows from the control plane to the data plane. It improves the communication efficiency and reliability of the whole system.

**ACKNOWLEDGMENT**

This study was supported in part by a grant from the National Basic Research Program of China (973 program) under Grant No. 2012CB315902, the National Natural Science Foundation of China under Grant No.60903214, 60970126.

**REFERENCES**

Abed, G.A., M. Ismail and K. Jumari, 2011. Behavior of cwnd for TCP source variants over parameters of LTE networks. *Inform. Technol. J.*, 10: 663-668.

Athuraliya, S., S.H. Low, V.H. Li and Q. Yin, 2001. REM: Active queue management. *IEEE Network*, 15: 48-53.

Audu, A., R. Gopal, H. Khosravi and C. Wu, 2003. Forwarding and control element protocol (FACT). <http://tools.ietf.org/id/draft-gopal-forces-fact-06.txt>

Aweya, J., 2001. IP router architectures: An overview. *Int. J. Commun. Syst.*, 14: 447-475.

Bennett, J.C.R. and H. Zhang, 1996. WF<sup>2</sup>Q: Worst-case fair weighted fair queueing. *Proc. IEEE INFOCOM*, 1: 120-128.

Bennett, J.C.R. and H. Zhang, 1997. Hierarchical packet fair queueing algorithms. *IEEE/ACM Trans. Networking*, 5: 675-689.

Crouch, A., H. Khosravi, A. Doria, X. Wang and K. Ogawa, 2010. ForCES applicability statement. RFC 6041. <http://www.ietf.org/proceedings/75/slides/forces-4.pdf>

Dong, L.G., L. Cai, L.F. Zhu and W.M. Wang, 2011. Research on high availability of ForCES control element. *Inform. Technol. J.*, 10: 1683-1691.

Doria, A., J.H. Salim, R. Haas, H. Khosravi and W.M. Wang *et al.*, 2010. Forwarding and control element separation (ForCES) protocol specification. RFC 5810. <http://tools.ietf.org/html/rfc5810>

Feng, W., D. Kandlur and D. Saha, 1999. A self-configuring RED gateway. *Proceedings of IEEE 18th Annual Joint Conference of the IEEE Computer and Communications Societies*, March 21, 1999, IEEE Computer Society, New York, USA., pp: 1320-1328.

- Floyd, S. and V. Jacobson, 1993. Random early detection gateways for congestion avoidance. *IEEE/ACM Transa. Networking*, 1: 397-413.
- Fu, J., 2008. On the design of next-generation routers and IP networks. Ph.D. Thesis, KTH, Stockholm, Sweden.
- Guo, C., 2001. SRR: An  $O(1)$  time complexity packet scheduler for flows in multi-service packet networks. *Proceedings of the ACM SIGCOMM*, Aug. 27-31, 2001, San Diego, California, USA., pp: 211-222.
- Haleplidis, E., K. Ogawa, W. Wang and J.H. Salim, 2010. Implementation report for ForCES. <http://datatracker.ietf.org/doc/draft-ietf-forces-implementation-report/>
- Halpern, J. and H. Ma, 2007. A VPN library for use with the ForCES protocol and model. <http://tools.ietf.org/html/draft-halpern-forces-lfblibrary-vpn-00>
- Halpern, J., 2005. A base library for use with the ForCES protocol and model. <http://tools.ietf.org/html/draft-halpern-forces-lfblibrary-base-00>.
- Hollot, C.V., V. Misra, D. Towsley and W.B. Gong, 2001. On designing improved controllers for AQM routers supporting TCP flows. *Proc. IEEE Commun. Comput. Soci.*, 3: 1726-1734.
- Ibrahim, A.A., 2008. Some transformation schemes involving the special (132)-avoiding permutation patterns and a binary coding: an algorithmic approach. *Asian J. Algebra*, 1: 10-14.
- Jasem, H.N., Z.A. Zukarnain, M. Othman and S. Subramaniam, 2010. The delay with new-additive increase multiplicative decrease congestion avoidance and control algorithm. *Inform. Technol. J.*, 4: 1327-1335.
- Kanher, S.S. and H. Sethu, 2001. Fair efficient and low-latency packet scheduling using nested deficit round robin. *Proceedings of the IEEE Workshop on High Performance Switching and Routing*, May 29-31, 2001, Dallas, TX., USA., pp: 6-10.
- Khosravi, H. and T. Anderson, 2003. Requirements for Separation of IP Control and Forwarding. RFC 3654, The Internet Society. <http://www.ietf.org/rfc/rfc3654.txt>
- Kunniyur, S. and R. Srikant, 2000. Analysis and design of an Adaptive Virtual Queue (AVQ) algorithm for active queue management. *Assoc. Comput. Machinery Comput. Commun. Rev.*, 31: 123-134.
- Matsurum, N. and R. Aibara, 1999. Efficient fair queueing for ATM networks using uniform round robin. *Proc. IEEE INFOCOM*, 1: 389-397.
- Mohammadi, S.M.A., A.A. Gharaveisi and M. Mashinchi, 2009. A novel fast and efficient evolutionary method for optimal design of proportional integral derivative controllers for automatic voltage regulator systems. *Asian J. Applied Sci.*, 2: 275-295.
- Ogawa, K., W.M. Wang, E. Haleplidis and J.H. Salim, 2011. ForCES Intra-NE high availability draft-ietf-forces-ceha-01. <http://tools.ietf.org/html/draft-ietf-forces-ceha-01>.
- Ott, T.J., T.V. Lakshman and L.H. Won, 1999. SRED: Stabilized RED. *Proceedings of the IEEE Infocom*, March, 1999, IEEE Computer Society, New York, USA., pp: 1346-1355.
- Qaddoura, E.A., B. El-Asir and M.M. Aqel, 2002. Enhanced aggressive error recovery for TCP Noor. *Inform. Technol. J.*, 1: 80-88.
- Salim, J.H., R. Haas and S. Blake, 2003. Netlink2 as ForCES Protocol. Internet-Draft, October 26, 2003, The Internet Society. <http://tools.ietf.org/id/draft-jhsrha-forces-netlink2-02.txt>
- Sharma, S., R.C. Jain and S.S. Bhadauria, 2007. SBEVA: A secured bandwidth efficient variance adaptive routing protocol for mobile Ad hoc network. *Asian J. Inform. Manage.*, 1: 1-10.
- Stoica, I., S. Shenker and H. Zhang, 2003. Core-stateless fair queueing: A scalable architecture to approximate fair bandwidth allocations in high-speed networks. *Inst. Electr. Electron. Eng. Assoc. Comput. Mach. Trans. Networking*, 11: 33-46.
- Takourout, R.N., S. Pierre and L. Marchand, 2006. Separation of the control plane and forwarding plane in next-generation routers. *J. Comput. Sci.*, 2: 815-823.
- Wang, W. and L. Dong, 2004. Design and performance evaluation of general router management protocol (GRMP). *Proc. IEEE Region 10th Conf. TENCN*, 2: 37-40.
- Wang, W.M.Y., F. Guo, G.M. Wang and L.G. Dong, 2004. General Router Management Protocol (GRMP) Version 1. Internet Draft, June 2004. <http://tools.ietf.org/id/draft-wang-forces-grmp-02.txt>
- Yang, L., R. Dantu, T. Anderson and R. Gopal, 2004. Forwarding and control element separation (ForCES) framework. RFC 3746. Network Working Group, The Internet Society. <http://www.ietf.org/computing/rfc/rfc3746.html>
- Zhang, Y., M. Roughan, W. Willinger and L. Qiu, 2009. Spatio-temporal compressive sensing and internet traffic matrices. *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*, August 16-21, 2009, ACM, New York, USA., pp: 267-278.
- Zhuge, B., C. Yu, K.P. Liu and W.M. Wang, 2011. Research on internal flow control mechanism of for CES routers. *Inform. Technol. J.*, 10: 626-638.