

QoS-enhanced TNPOSS Network Model and its E2E Delay Bound for Multimedia Flows

Ke Xiong

Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
E-mail: xiongke.bjtu.iis@gmail.com

Yu Zhang, Zheng-ding Qiu, Dong-mei Sun

Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China

Ya-Juan Qin

School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

Abstract—In our previous work, we proposed the TNPOSS (To Next-hop Port Sequence Switch) network which can achieve scalable fast forwarding and is suitable for delivering multimedia flows. However, TNPOSS network has no Quality of Service (QoS) tools to provide better QoS guarantee for multimedia flows which often have long-range dependence (LRD) property. To enhance the QoS ability of TNPOSS network, this paper proposes a QoS-enhanced TNPOSS (QTNPOSS) network model by introducing Fractal Leak Bucket (FLB) shaper and Weighted Fair Queuing (WFQ) scheduler into the original TNPOSS network. In the new QTNPOSS network, each packet of multimedia flows is shaped by the FLB when arriving at a node and is scheduled by the WFQ before being outputted. Based on study above, we further analyze the end-to-end (E2E) delay bound of QTNPOSS network by means of the network calculus theory which is an effective mathematical tool on analyzing the worst-case QoS performances of networks. The service curve and the formulation of E2E delay bound of QTNPOSS network are presented and proved. Extensive numerical experiments show that both the LRD property of multimedia flows and the WFQ weight have influences on the E2E delay bound, and the WFQ weight influences the E2E delay bound more greatly than the LRD property does.

Index Terms—QTNPOSS network, end-to-end delay bound, multimedia flow, TNPOSS network, fractal leak bucket, WFQ, network calculus

I. INTRODUCTION

A. Acronyms & Notations

BDF Burst-delay Function

E2E	End to End
FLB	Fractal Leak Bucket
GPS	Generalized Processing Sharing
LRD	Long Range Dependence
QoS	Quality of Service
QTNPOSS	QoS-enhanced TNPOSS
RLF	Rate-Latency Function
SS	Self-Similar
TB	Token Bucket
TNPOSS	To Next-hop Port Sequence Switch
WFQ	Weighted Fair Queuing
WIF	Wide-sense Increasing Function
δ_T	BDF with delay T
$\beta_{V,T}$	RLF with rate V and latency T
ξ	Node processing delay
φ	Link propagation delay

B. Background & Motivation

With the rapid development of network technology, more and more multimedia applications are expected to be delivered over IP-based Internet. Such applications often have strict QoS requirements on network metrics, including end to end (E2E) delay, jitter and packet loss probability, etc.

To meet the transmission requirements of multimedia flows, we have proposed the TNPOSS forwarding approach in the previous work [1], which can perform scalable fast forwarding to achieve lower E2E delay. Since TNPOSS network adopts connection-oriented forwarding mechanism and works like explicit routing, it is capable of improving the QoS ability of packet networks. However, no other QoS tools have been designed for TNPOSS network to further enhance its QoS performances.

To enhance the QoS ability of TNPOSS network, in this paper, we add two important components into the original TNPOSS network and then propose a novel QoS-enhanced TNPOSS (QTNPOSS) network model. One of the components added by us is the traffic shaping module and the other is the queue scheduling block. One on hand,

This paper is supported partially by the National Basic Research Program of China (“973 program”) under contract No. 2007CB307101 and No. 2007CB307106, partially by the National High-Tech Research and Development Plan of China under Grant No. 2007AA01Z202, partially by the Program of Introducing Talents of Discipline to Universities (“111 Project”) under contract No. B08002, partially by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China under contract No. 706005, and partially by the National Natural Science Funds under grant No. 60772043.

Corresponding author: Ke Xiong (Beijing Jiaotong University, China, xiongke.bjtu.iis@gmail.com)

because most internet flows have very high and stochastic burst rate, they must be shaped upon arriving at each node [2][3]. As is known, multimedia traffic often has the self-similar (SS) or long-range dependence (LRD) property. Thus, we introduce the fractal leak bucket (FLB) [4] model into our QTNPOSS network to perform traffic shaping for multimedia flows. On the other hand, there may be often several flows going through a same network node at a same time, so queue scheduling device is required to provide differentiated service for such flows. Among existing queue scheduling mechanisms, weighted fairness queuing (WFQ) [5] model is considered as one of the most effective ones, so we introduce WFQ into the QTNPOSS network.

In network performance metrics, E2E delay is one of the most important targets of QoS guarantee and its bound has great significance for network congestion control, buffer-size adjustment and scheduling optimization. Therefore, we analyze the E2E delay bound of QTNPOSS network in this paper. The analysis tool we use is the Network Calculus [6], which is an effective mathematical tool on analyzing network performances quantitatively.

C. Contributions & Organization

The main contributions and novelties of this paper are: 1) proposing the QTNPOSS network model to improve the QoS ability of TNPOSS network; 2) modeling the E2E delay of the QTNPOSS network; 3) giving the formulation of the E2E delay bound of QTNPOSS network; 4) analyzing the parameters' influences on E2E delay bound of the QTNPOSS network by numerical experiments.

The rest of this paper is organized as follows: Section II is the related work. The original TNPOSS network, the traffic shaping models, the WFQ scheduler and the tools of network calculus are introduced. In section III, we give the network model of the QTNPOSS network. Based on these, in Section IV, we analyze the E2E delay bound of QTNPOSS network. In Section V, extensive numerical experiments are performed to discuss the parameters' influences on the E2E delay bound. Finally, Section VI summarizes the work of this paper with some concluding remarks.

II. RELATED WORK

A. TNPOSS Network

In TNPOSS Network, a set of binary codes are used to identify the ports of a router. For example, consider the network shown in Figure 1, S and D are terminal devices. a , b , c , d and e represent the routing devices. The solid lines between any two nodes are the communication links. The binary code near each node is the port code of the corresponding link interface. Suppose $S \rightarrow a \rightarrow b \rightarrow d \rightarrow c \rightarrow e \rightarrow D$ is the selected path for delivering the packets from S to D , the path can be represented by the output port sequence 10 11 11 100 01 which consists of the port code of the output link at each hop on the path.

If the sender S wants to send multimedia flows to its

destination D , it will initiate a PATH_SETUP request to setup one or more output-port code sequence paths at first. Then it can use the output port sequence to do explicit routing.

In TNPOSS network, just during the path establishing stage, routing tables have to be visited, but during the data sending stage, each packet is forwarded directly according to the path code in the packet's header, so the cost of routing lookup is greatly reduced and it can achieve fast switching. Moreover, TNPOSS is able to work on any routing protocol. If the routing protocol provides QoS routing, TNPOSS will transmit the packets on QoS-supported path. Since TNPOSS is able to deliver the packets of a flow on the specified connection-based path, it has better QoS ability than today's IP networks. The detailed working procedure of TNPOSS network can be found in [1].

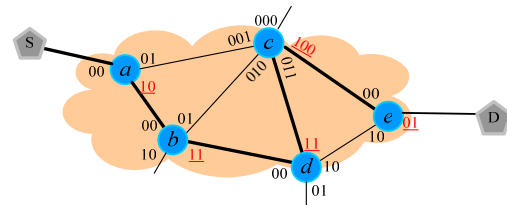


Figure. 1 TNPOSS network model

Although TNPOSS is suitable for transmitting multimedia flows due to its special working procedure and its explicit routing nature, it has no QoS tools to meet the strict QoS requirements of multimedia applications. Thus, this paper aims to improve the QoS performances of TNPOSS network.

B. Fractal Leaky Bucket (FLB)

To enhance network QoS ability, traffic is required to be policed in order to guarantee that the sender does not send more than specified by the contract of a connection into the network. Policing devices inside network often do buffering, which are called shapers. One of the simple and effective shaper models is token bucket (TB) [4]. TB regulates the traffic by a linear function of a time interval τ .

If we denote the traffic the sender transmits over the time interval τ with $A(\tau)$, the traffic is said to be regulated by TB, if there exists a pair (ρ, b) such that:

$$A(\tau) \leq \rho\tau + b \text{ for any } \tau > 0. \quad (1)$$

Where ρ represents the long-term average rate of the traffic (i.e. the output rate of the TB) and b represents the maximum burst allowed to be sent into the network in any time interval τ (i.e. the buffer size of the TB). TB has good ability to describe the characteristic of the linear-bounded arrival processor, but it can not describe the LRD traffic very well. More detailed information about TB can be found in [7].

[2][3][4] state that the LRD property may bring down network performances, including increasing the E2E delay, the buffer size and the packet loss probability, etc. Thus, it is necessary to select a proper traffic shaper for TNPOSS network to support multimedia traffic. [4] proposes the FLB instead of TB to regulate the LRD traffic and the

numerical results show that FLB outperforms TB in terms of shaping LRD flows. Hence, we choose the FLB shaper for TNPOSS network.

If we denote the traffic the sender transmits over the time interval τ with $A^*(\tau)$, the traffic is said to be regulated by FLB, if there exists a pair (ρ^*, b^*) such that:

$$A^*(\tau) \leq \rho^* \tau + b^* \text{ for any } \tau > 0. \quad (2)$$

Where

$$\rho^* = \rho + \sigma(1-H) \sqrt{2\gamma \left(\frac{H}{1-H}\right)^{H-1}} \quad (3)$$

and

$$b^* = \sigma(1-H) \sqrt{2\gamma \left(\frac{H}{1-H}\right)^H}. \quad (4)$$

Where γ is a positive constant, whose value is usually assumed to be 6 [10]. σ is the standard variation of $A(\tau)$, and H is the self-similar parameter, which is in fact the burst parameter.

C. Weighted Fair Queuing (WFQ)

Scheduling is one of the most important mechanisms to provide QoS guarantee in packet networks. One of the most notable scheduling models is the Generalized Processing Sharing (GPS) model [5][6]. GPS can control the sharing of one link among packets of different classes, but it is only an ideal method and is not implementable. To approximate GPS, WFQ is considered as the most effective one, which does not make the assumption of infinitesimal packet size. So, we introduce WFQ scheduler into QTNOSS network.

Suppose a WFQ scheduler serves N flows and each flow is specified by a positive weight w_i . $g_i(\tau)$ denotes the amount of served traffic of flow i in the time interval τ . If R is the service rate of the network node,

$$g_i(\tau) = \frac{w_i}{\sum_{1 \leq k \leq N} w_k} * R * \tau. \quad (5)$$

D. Network Calculus

Network calculus [6] is a collection of results based on Min-Plus algebra, which can be applied to deterministic queuing systems found in communication networks. It is a set of recent developments which provide a deep insight into flow problems encountered in networking, and is used with envelope bounded traffic models to provide a worst-case analysis on network performance. Network calculus is based on the idea that given a regulated flow of traffic into the network, one can quantify the characteristics of the flow as it travels from node to node through the network. This means that traffic flows are constrained by shapers and delayed by the nodes' schedulers.

In network calculus, the shapers are modeled by arrival curves and the schedulers are modeled by latency service curves. Now we introduce some important tools and conclusions of network calculus as follows.

Definition 1. (WIF: wide-sense increasing function). $f(x)$ is a function, for any $s \leq t$, if $f(s) \leq f(t)$, f is a wide-sense increasing function.

The WIF is used to describe flow functions such as the

$A(\tau)$ and $A^*(\tau)$ in this Section.

Definition 2. (arrival curve). Give a WIF α defined for a shaper, a flow f is constrained by α if and only if for all $s \leq t$,

$$f(s) - f(t) \leq \alpha(t - s). \quad (6)$$

The arrival curve actually defines an upper bound on the arrival rate of a flow to a particular node. The arrival curve of FLB is can be modeled by Equation (2).

Definition 3. (service curve). If a system S has an input flow $f(t)$ and output flow $f_o(t)$, then S offers to the flow a service curve $\beta(t)$, if and only if for all $t \geq 0$,

$$f_o(t) \geq \inf_{s \leq t} (f(s) + \beta(t - s)). \quad (7)$$

A service curve is a lower bound on the departure rate from a network node.

Definition 4. (min-plus convolution). Let f and g be two WIFs. The min-plus convolution of f and g is the function:

$$(f \otimes g)(t) = \begin{cases} \inf_{0 \leq s \leq t} [f(t - s) + g(s)], & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (8)$$

Definition 5. (min-plus deconvolution). Let f and g be two WIFs. The min-plus deconvolution of f and g is the function:

$$(f \oslash g)(t) = \sup_{s \geq 0} [f(t + s) - g(s)] \quad (9)$$

Definition 6. (virtual delay). The virtual delay at time t is

$$d(t) = \inf\{\tau \geq 0: f(t) \leq f_o(t + \tau)\} \quad (10)$$

The virtual delay at time t is the delay that would be experienced by a bit arriving at time t if all bits received before it are served before it.

Definition 7. (backlog). If a system S has an input flow $f(t)$ and output flow $f_o(t)$, the backlog at time t is defined as

$$\phi(t) = f_o(t) - f(t) \quad (11)$$

The backlog is the amount of bits that are held inside the system; if the system is a single buffer, it is the queue length.

Definition 8. (horizontal deviation). Let f and g be two WIFs. The horizontal deviation $h(f, g)$ is defined as

$$h(f, g) = \sup_{t \geq 0} \{\inf\{\tau \geq 0; f(t) \leq g(t + \tau)\}\} \quad (12)$$

Definition 9. (vertical deviation). Let f and g be two WIFs. The vertical deviation $v(f, g)$ is defined as

$$v(f, g) = \sup_{t \geq 0} \{f(t) - g(t)\} \quad (13)$$

Definition 10. (BDF: burst-delay function). WIF $\delta_r(t)$ is called burst-delay function, if

$$\delta_r(t) = \begin{cases} 0, & t < T \\ +\infty, & t \geq T \end{cases} \quad (14)$$

Definition 11. (RLF: rate-latency function). WIF $\beta_{v,T}(t)$ is called rate-latency function, if

$$\beta_{v,T}(t) = V[t - T_{lat}]^+ = \begin{cases} V(t - T_{lat}) & t > 0 \\ 0 & t \leq 0 \end{cases} \quad (15)$$

Where T_{lat} is the latency delay and V is the service rate. The service curve of a GPS node can be represented by a RLF [6].

Property 1. (service curve of concatenation nodes). Assume a flow traverses systems $S_1, S_2, \dots,$ and S_m in sequence. Suppose that $S_i, i \in [1, m],$ offers a service curve of β_i to the flow. Then the concatenation of the systems offers a service curve of $\beta_1 \otimes \beta_2 \otimes \dots \otimes \beta_m$ to the flow.

Property 2. $d(t) \leq h(\alpha, \beta).$

Property 2 shows that the virtual delay of a system is the horizontal deviation between its arrival curve α and service curve $\beta.$

Property 3. $\phi(t) \leq v(\alpha, \beta).$

Property 3 shows that the backlog of a system is the vertical deviation between its arrival curve α and service curve $\beta.$

Property 4. $\delta_{T_1} \otimes \delta_{T_2} = \delta_{T_1+T_2}.$

Property 5. $\beta_{V, T_1} \otimes \delta_{T_2} = \beta_{V, T_1+T_2}.$

Property 6. $\beta_{V_1, T_1} \otimes \beta_{V_2, T_2} = \beta_{\min\{V_1, V_2\}, (T_1+T_2)}.$

Given three WIFs f, g and $\zeta,$ the following properties hold in network calculus.

Property 7. $f \otimes g = g \otimes f.$

Property 8. $f \otimes g \otimes \zeta = f \otimes (g \otimes \zeta).$

More detailed information can be found in [6].

Recently, many researches [8][9][10] analyze network QoS performances with network calculus. [8] discusses the E2E delay bound of the Expedited Flow defined in RFC 3246. [9] analyzes the E2E delay bound of the wireless sensor networks via the statistical network calculus and [10] proposed the E2E delay bound for LRD flows under the shaping model of FLB. In this paper, we use network calculus to analyze the E2E delay bound of our QTNPOSS network. The detailed work will be described in Section IV.

III. QOS-ENHANCED TNPOSS NETWORK

From the description of Section II.A, it can be seen that TNPOSS network performs QoS provisioning just via explicit routing mechanism and the output-port-code based fast forwarding. No additional tools are provided to enhance the QoS ability of TNPOSS network.

In this section, we add FLB shapers and WFQ schedules into the nodes of TNPOSS networks and call the new network model the QoS-enhanced TNPOSS (QTNPOSS) Network.

Figure 2 gives a topology example and the node structure of QTNPOSS Network.

When a packet of a flow arrives at a node of QTNPOSS network, it is shaped by the FLB shaper, and then is scheduled by the WFQ scheduler according to the weight of the flow it belongs to.

Suppose the service rate of a QTNPOSS node is $R,$ in terms of Equation (5), the service rate V_i for flow i is

$$V_i(t) = \frac{w_i}{\sum_{1 \leq k \leq N} w_k} * R(t) \tag{16}$$

Moreover, according to [11], the maximum delay of WFQ model is

$$d_{wfq_max}^i = \frac{l_{max}^i}{V_i} + \frac{L_{max}}{R} \tag{17}$$

Where l_{max}^i is the maximal packet length of flow i and L_{max} is the maximal packet length of all flows in the node.

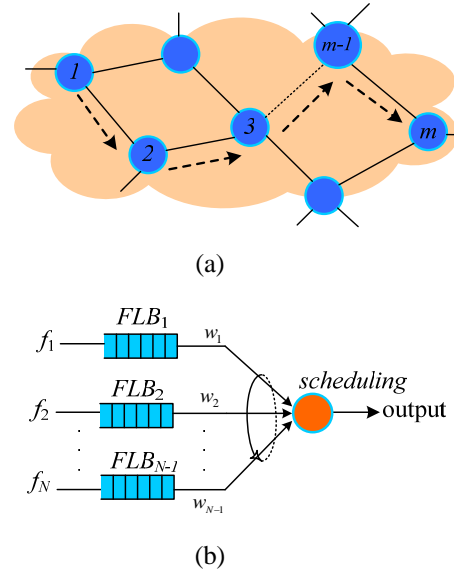


Figure 2 QTNPOSS network (a) and its node structure (b)

IV. END TO END DELAY BOUND

The components of the E2E delay of QTNPOSS network can be divided into two classes. The first class is denoted by D_v consisting of the shaping delay and the scheduling delay, and the second class is denoted by D_s consisting of the link propagation delay and the node processing delay. The delay of the first class is variable, while the delay of the second class is relatively stable.

Theorem 1. (single node QTNPOSS network E2E delay bound). Suppose a LRD flow goes through a QTNPOSS network consisting of only one node. The delay caused by this node satisfies

$$d_{E2E} \leq \frac{b_i + l_{max}^i}{V_i} + \frac{L_{max}}{R} + \xi \tag{18}$$

Where, ξ denotes the forwarding processing time and the other parameters have the same definitions with those of the corresponding parameters in Equation (17).

Proof: Since the single node has the arrival curve of FLB and the service curve of WFQ, we denote them in a time-bit coordinate plane of Figure 3.

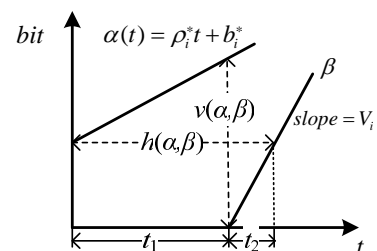


Figure 3 Arrival curve and service curve of a single node QTNPOSS network for flow i

From Figure 3 and Property 2, it can be inferred that

$$D_v \leq h(\alpha, \beta) = t_1 + t_2 \quad (19)$$

According to Equation (17),

$$t_1 = \frac{l_{\max}^i}{V_i} + \frac{L_{\max}}{R} \quad (20)$$

Hence,

$$\begin{aligned} d_{E2E} &\leq D_{\max}^s \leq D_v + D_s = t_1 + t_2 + D_s \\ &= \frac{l_{\max}^i}{V_i} + \frac{L_{\max}}{R} + \frac{b_i}{V_i} + D_s \end{aligned} \quad (21)$$

Since in the single node system, there is no link propagation delay, thus D_s only consists of ξ . So,

$$d_{E2E} \leq \frac{b_i + l_{\max}^i}{V_i} + \frac{L_{\max}}{R} + \xi \quad \square$$

Theorem 2. (single node QTNPOSS network backlog).

Suppose a LRD flow goes through a QTNPOSS network consisting of only one node. The backlog of the network must satisfies

$$\phi_s \leq \rho_i \left(\frac{l_{\max}^i}{V_i} + \frac{L_{\max}}{R} \right) + b_i \quad (22)$$

Proof: According to the definition 7, property 3 and Figure 3, it is easily to be inferred that

$$\phi_s = v(\alpha, \beta) \leq \rho_i^* t_1 + b_i \quad (23)$$

By substitution of the Equation (20) into (23), we can prove that

$$\phi_s \leq \rho_i \left(\frac{l_{\max}^i}{V_i} + \frac{L_{\max}}{R} \right) + b_i \quad \square$$

Theorem 3. (QTNPOSS network service curve).

Assume that a LRD flow goes through a QTNPOSS network and the service curves of the m nodes on the transmission path are β_{V_1, T_1} , β_{V_2, T_2} , β_{V_3, T_3} , ..., β_{V_m, T_m} , respectively. Denote the propagation delays of the links among the m nodes with $\varphi_1, \varphi_2, \dots, \varphi_{m-1}$ in turn, and denote the node processing delay of the m nodes with $\xi_1, \xi_2, \xi_3, \dots, \xi_m$, respectively. The service curve of the QTNPOSS network is

$$\beta_{E2E} = \beta_{\min\{V_1, V_2, \dots, V_m\}, \left(\sum_{1 \leq i \leq m} T_i + \sum_{1 \leq i \leq m} \xi_i + \sum_{1 \leq i \leq m-1} \varphi_i \right)} \quad (24)$$

Proof: As both the link propagation service curve and the node processing service curve can be represented by BDF function [6], we denote their service curves with δ_{ξ_k} and δ_{φ_k} respectively. Now, we prove this theorem by the mathematical induction.

1) When $m = 1$, $\beta_{E2E} = \beta_{V_1, T_1} \otimes \delta_{\xi_1}$. According to

Property 5, $\beta_{E2E} = \beta_{V_1, T_1 + \xi_1}$. (25)

2) When $m = k - 1$, suppose that

$$\hat{\beta}_{E2E} = \beta_{\min\{V_1, V_2, \dots, V_{k-1}\}, \left(\sum_{1 \leq i \leq k-1} T_i + \sum_{1 \leq i \leq k-2} \xi_i \right)} \quad (26)$$

holds.

3) When $m = k$, according to Property 1,

$$\beta_{E2E} = \hat{\beta}_{E2E} \otimes \beta_{V_k, T_k} \otimes \delta_{\xi_k} \otimes \delta_{\varphi_{k-1}} \quad (27)$$

Based on Property 7, Property 5, Property 6 and Property 8, it can be obtained that

$$\beta_{E2E} = \beta_{\min\{V_1, V_2, \dots, V_k\}, \left(\sum_{1 \leq i \leq k} T_i + \sum_{1 \leq i \leq k-1} \xi_i \right)} \quad (28)$$

□

Theorem 4. (QTNPOSS network E2E delay bound).

Assume that a LRD flow goes through a QTNPOSS network, the arrival curves of the m nodes on the transmission path are $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_m$ respectively for all $1 \leq i \leq m$, where $\alpha_i = \rho_i^* t + b_i^*$. The service curves of the m nodes are $\beta_{V_1, T_1}, \beta_{V_2, T_2}, \dots, \beta_{V_m, T_m}$. Let the propagation delays of the links among the m nodes be $\varphi_1, \varphi_2, \dots, \varphi_{m-1}$ in turn and the node processing delay of the m nodes be $\xi_1, \xi_2, \xi_3, \dots, \xi_m$, respectively. The E2E delay D_{E2E} of the QTNPOSS network satisfies that

$$D_{E2E} \leq \sum_{1 \leq k \leq m} T_k + \frac{b_i^*}{\min\{V_1, V_2, \dots, V_m\}} + \sum_{1 \leq k \leq m} \xi_k + \sum_{1 \leq k \leq m-1} \varphi_k \quad (29)$$

Where $T_k = d_{WFQ_max}^k$, which denotes the maximal delay of the WFQ scheduler in the k -th node of the transmission path and V_k is the service rate of the flow in the k -th node. $d_{WFQ_max}^k$ is calculated by Equation (17) and V_k is calculated by Equation (16).

Proof: Under assumption of Theorem 3, the whole QTNPOSS network can be seen as a virtual node with the arrival curve of α_1 in the network edge and the service curve of Equation (24). The arrival curve and the service curve of such virtual node are shown in Figure 4.

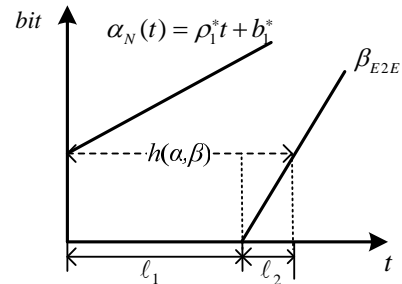


Figure 4 Arrival curve and service curve of QTNPOSS network

According to property 2,

$$D_{E2E} \leq h(\alpha_N, \beta_{E2E}) = l_1 + l_2.$$

Based on Definition 11 and Equation (24),

$$l_1 = \sum_{1 \leq k \leq m} d_{WFQ_max}^k + \sum_{1 \leq k \leq m} \xi_k + \sum_{1 \leq k \leq m-1} \varphi_k \quad (30)$$

$$l_2 = \frac{b_i^*}{\min\{V_1, V_2, \dots, V_m\}} \quad (31)$$

Therefore,

$$\begin{aligned} D_{E2E} &\leq l_1 + l_2 = \sum_{1 \leq k \leq m} d_{WFQ_max}^k + \sum_{1 \leq k \leq m-1} \xi_k \\ &\quad + \sum_{1 \leq k \leq m} \varphi_k + \frac{b_i^*}{\min\{V_1, V_2, \dots, V_m\}} \end{aligned}$$

□

Corollary 1. When $m = 1$, the conclusion of Theorem 4 is identical to that of Theorem 1.

Proof: This corollary can be proved easily by assuming $m = 1$ in Equation (29). □

Corollary 2. The greedy FLB does not cause extra increase of the E2E delay of QTNPOSS.

Proof: From Equation (29) and Figure 4, we can see that the $\alpha_2, \alpha_3, \dots, \alpha_m$ have no influence on E2E delay of QTNPOSS network. □

A similar conclusion of corollary 2 was proposed in [10], which is about LRD flows.

V. NUMERICAL ANALYSIS

In this Section, we analyze the parameters' influences on E2E delay bound of QTNPOSS network. The parameters, including the self-similar parameter H , the WFQ weight w and the number of nodes m are taken into consideration in our analysis. The other parameters and their values used in the numerical experiments of this section are listed in the Table I.

TABLE I.
THE PARAMETERS AND THEIR VALUES

name	value	unit	name	value	unit
ρ	4	kbit/s	ξ	0.1	ms
σ	8	kbit	φ	3×10^{-4}	kbit
γ	6	-	R	10	Mbit/s
l_{\max}	1280	byte	L_{\max}	1280	byte

Where ρ , σ and γ are defined in Equation (3) and (4). R is the service rate of the node in QTNPOSS network. ξ and φ are the parameters defined in Equation (24). l_{\max} and L_{\max} are used in Equation (17).

Since we aim to analyze the E2E delay bound, which is the E2E delay in the worst case, we assume that all the nodes and links have the same processing delay ξ and the same propagation delay φ respectively in the numerical experiments. Moreover, l_{\max} and L_{\max} are assumed to be the maximal value of the MTU (maximum transmission unit), which is about 1280 bytes.

A. E2E delay of single node QTNPOSS network

In this Section, the QTNPOSS network is assumed to have only one node, which was introduced in Theorem 1. Figure 5 and Figure 6 show the influences of the self-similar parameter H and the WFQ weight w on the E2E delay bound of the single node QTNPOSS network, respectively.

From Figure 5 and Figure 6, we can see that the E2E delay bound of single node QTNPOSS network decreases with the increase of both H and w . The smaller the H and w are, the higher the E2E delay bound is. Moreover, the results indicate that the influence of w on the E2E delay bound is bigger than that of H , especially when w is small. Figure 6 shows that when $w < 0.3$, the E2E delay

increases very fast with the decrease of w .

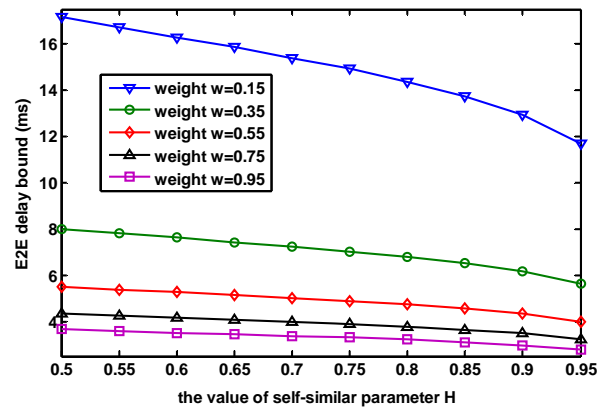


Figure 5 E2E delay bound versus H in single node QTNPOSS network

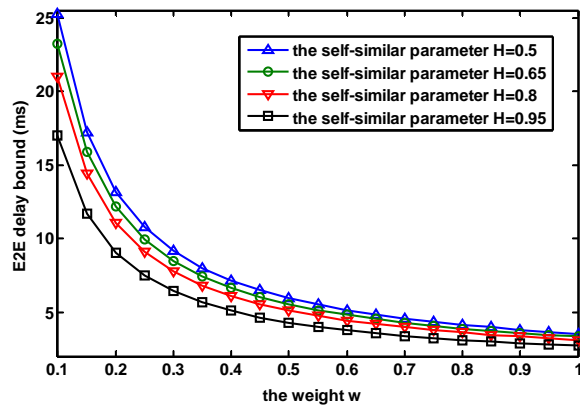


Figure 6 E2E delay bound versus w in single node QTNPOSS network

B. E2E delay of multi-node QTNPOSS network

In this section, a flow is assumed to go through a path consisting of m nodes in the QTNPOSS network. In the experiments of Figure 7 and Figure 8, m is assumed to be 15, because the measurement work in [12][13] shows that the average number of hops between two communication nodes in Internet is about 15. Figure 7 shows the influence of the self-similar parameter H on the E2E delay bound of and Figure 8 shows the influence of the WFQ weight w on the E2E delay bound.

From Figure 7 and Figure 8, it also can be inferred that the E2E delay bound of Q TNPOSS network is influenced by both H and w and the influence of w is greater than that of H , especially when w is small.

In the experiments of Figure 9, m ranges from 10 to 25. The results in Figure 9 shows that the E2E delay bound grows with the increasing of m and the growing rate shows linear growth trend.

From all the experiments above, we can state that w influences the E2E delay bound more greatly than the other parameters do. The numerical results also indicate that by means of rising the degree of a flow's self-similar property or elevating the flow's WFQ weight, one can achieve acceptable low E2E delay.

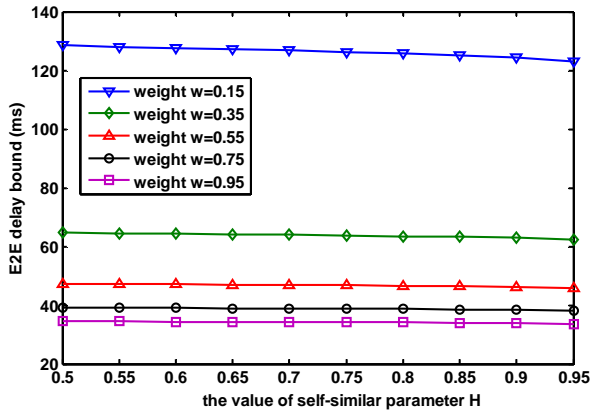


Figure 7 E2E delay bound versus H in QTNPOSS network with $m > 1$ nodes

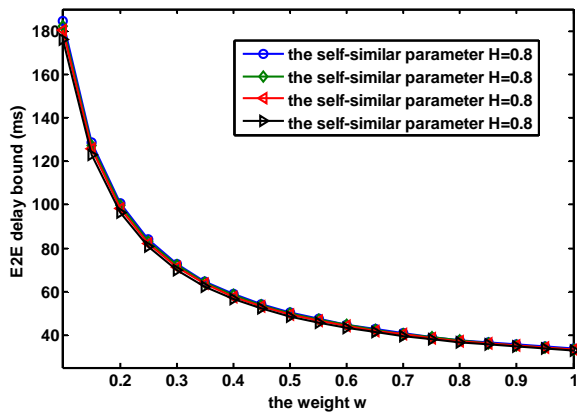


Figure 8 E2E delay bound versus w in QTNPOSS network with $m > 1$ nodes

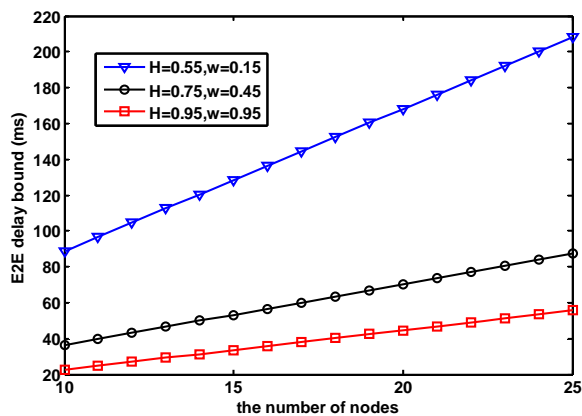


Figure 9 E2E delay bound versus m in QTNPOSS network

VI. CONCLUSIONS

This paper proposes the QTNPOSS network model by introducing the FLB shaper and the WFQ scheduler into TNPOSS network. Based on this, we analyze the E2E delay bound of the QTNPOSS network using network calculus. We give the service curve and the formulation of E2E delay bound of QTNPOSS network. Extensive

numerical experiments show that both the long-range dependence property and the WFQ weight have influence on the E2E delay bound, and the WFQ weight has greater influence on the E2E delay bound than that of the long-range dependence property.

ACKNOWLEDGEMENTS

The authors wish to give many thanks to Prof. Hong-ke Zhang, Prof. Yu-chun Guo and Dr. Gao, all of whom are with the School of Electronic and Information Engineering, Beijing Jiaotong University, China, because they help us so much for this paper. This paper is supported partially by the National Basic Research Program of China ("973 program") under contract No. 2007CB307101 and No. 2007CB307106, partially by the National High-Tech Research and Development Plan of China under Grant No. 2007AA01Z202, partially by the Program of Introducing Talents of Discipline to Universities ("111 Project") under contract No. B08002, partially by the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China under contract No. 706005, and partially by the National Natural Science Funds under grant No. 60772043.

REFERENCES

- [1] K. Xiong, Z. D. Qiu, H. K. Zhang, C. S. Li, H. C. Chao, "A scalable fast forwarding approach for IP networks," *International Journal of Internet Protocol Technology*, vol. 3, no. 2, 2008, pp. 119-127.
- [2] J. Y. Le Boudec, "Some Properties Of Variable Length Packet Shapers," *ACM/IEEE Transactions on Networking*, vol. 10, no. 3, 2002, pp. 329-337.
- [3] D. Chakraborty, A. Ashir, T. Sukanuma, G. Mansfield Keeni, T. K. Roy, N. Shiratori, "Self-similar and fractal nature of internet traffic," *International Journal of Network Management*, vol. 14, no. 2, 2004, pp. 119-129.
- [4] G. Procissi, A. Garg, M. Gerla and M. Y. Sandadidi, "Token bucket characterization of long-range dependent traffic," *Computer Communications*, vol. 25, no. 11-12, 2002, pp. 1009-1017.
- [5] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated service networks: The single node case," *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, 1993, pp. 344-357.
- [6] J. Y. Le Boudec and P. Thiran, *Network Calculus*, Berlin: Springer Verlag, Version May 10, 2004.
- [7] S. Keshav, *An engineering approach to computer networking: ATM networks, the Internet and the telephone network*, Addison-Wesley, 1997.
- [8] Y. M. Jiang, "Delay bound and packet scale rate guarantee for some expedited forwarding networks," *Computer Networks*, vol. 50, no. 1, 2006, pp. 15-28.
- [9] H. M. Qi, Z. Q. Chen, L. M. Zhang, "Towards end-to-end delay bounds on WMNs based on statistical network calculus," in *Proceedings of the 9th International Conference for Young Computer Scientists*, pp. 493-497, Nov. 2008.
- [10] L. M. Zhang, D. Z. Huang, Z. G. Chen, "Models of bounds on end-to-end delay of long-range dependence traffic based on fractal leak bucket," *Journal on Communications*, vol. 29, no. 7, 2008, pp. 32-43.

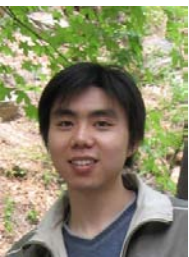
- [11] N. J. Towela, *Impact of Traffic Aggregation on Network Capacity and Quality of Service*, University of Kansas. <http://www.ittc.ukans.edu>.
- [12] J. G. Ma, M. X. Xi, Y. M. Lin, Y. P. Li, "Chinese Internet routerlevel hop count measurement and analysis," *Application Research on Computer*, vol. 25, no. 27, 2008, pp. 2112-2114.
- [13] F. Begtašević, P. Van Mieghem, "Measurements of the Hopcount in Internet," in *Proceedings of Passive and Active Measurement (PAM 2001)*, Amsterdam, the Netherlands, April 23-24, 2001, pp. 183-190.



Ke Xiong received his B.S. degree in computer science from Beijing Jiaotong University (formerly known as Northern Jiaotong University), Beijing, China, in June 2004. He is currently working toward the Ph.D. degree at the Institute of Information Science, School of Computer and Information Technology, Beijing

Jiaotong University, Beijing, China.

His research interests include signal processing, multimedia communication, next generation networks, network performances analysis, network resource management and QoS guarantee in IP networks.



Yu Zhang received his B.S. degree in computer science from Beijing Jiaotong University (formerly known as Northern Jiaotong University), Beijing, China, in June 2004. He is currently working toward the Ph.D. degree at the Institute of Information Science, School of Computer and Information Technology, Beijing

Jiaotong University, Beijing, China.

His research interests include digital signal processing, multimedia communication & processing and compressed sensing.



Zheng-ding Qiu received B.S. degree in Communication Engineering in 1967, and M.S. degree in Digital Signal Processing in 1981 from Northern Jiaotong University, Beijing respectively.

During 1967 to 1978, he was a communication engineer at LiuZhou Railway Bureau. Since 1981, he worked at Institute of Information Science, Northern Jiaotong University, where he became a lecturer in 1983, an associate professor in 1987 and a full professor in 1991. He was an overseas research scholar at Computer Science Department of Pittsburgh University, USA, and joined in the program of parallel processing in signal processing. During 1999 to 2000, he was a visiting

research fellow at Electronic Engineering Lab of Kent University, UK, and engaged in image processing and biometrics processing program.

From 1993 to 2001, professor Qiu undertook program "Research and Implement of Multimedia Conferences Terminal and System" which is one of Chinese Hi-tech projects and program "the Research and Implement of Multimedia Service Platform over IP network" which is a sub-project of Chinese National Ninth Five-year key projects.

His research field includes digital signal processing, multimedia communication and processing and parallel processing. He is currently a fellowship of Chinese Institute of Communications, a senior member of Chinese Institute of Electronics and Railway respectively.



Dong-mei Sun received her Ph.D. degree in Digital Signal Processing from Beijing Jiaotong University in 2003. She is now an associate professor at Institute of Information Science, Beijing Jiaotong University, China.

Currently Mrs. Sun has been doing some projects supported by National Natural Science Foundation of China, Specialized Research Fund for the Doctoral Program of Higher Education and National Science & Technology Pillar Program. Recently she has been cooperating with Imaging Research Center at University of Cincinnati (USA) on several medical image processing and analysis projects. So far she has published over 40 research papers in the domestic and international scientific journals and conferences.

Her research interests include biometric technology, image analysis, pattern recognition, information security and signal processing.



Ya-juan Qin received her B.S. and M.S. degrees in Radio Technology from the University of Electronic Science and Technology of China (formerly known as Chengdu Institute of Radio Engineering) in 1985 and 1988, respectively, and Ph.D. degree in communication engineering from Beijing University

of Posts and Telecommunications in 2003. From January 2002 to April 2002, she was a research associate at CRL, Japan. In 2004, she joined School of Electronics and Information Engineering, Beijing Jiaotong University (formerly known as Northern Jiaotong University), where she is now an associate professor. She has published more than 40 research papers.

Her research interests are in the areas of computer networks, next generation Internet and wireless communications.