TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.6, June 2014, pp. 4400 ~ 4404
DOI: 10.11591/telkomnika.v12i6.5474

Research on End-to-End Delay Performance Based on GPS Scheduling System

Wang Yang*, **Zhou Jin-he**, **Ye Keng**
School of Information and Telecommunication Engineering,
Beijing Information Science and Technology University,
Beijing 100101, China
*****Corresponding author, e-mail: wangyang282@126.com

Abstract

This paper summarizes some results of network calculus. It derives the bounds on end-to-end delay and effective bandwidth by network based on leaky bucket regulators. The simulation results show that the network calculus is a good fit for calculating end-to-end delay. And it provides effective control, scheduling and management in the integration network conditions in the following step.

Keyword: *statistical network calculus, GPS scheduling system, delay, effective bandwidth*

1. Introduction

Generalized Processor Sharing (GPS) is a kind of ideal type, continuous work of the fair scheduling strategy, put forward by Parekh and Gall anger [1]; because GPS provides a method in response to different service requests. So GPS becomes a research hotspot in the guaranteed service [2] network. Network calculus, a kind of new network system performance quantitative analysis, is important and effective as a teaching tool. Compared with the traditional method of queuing theory [3], the biggest advantage of network calculus is that it can provide network queue with deterministic analysis [4]. In this paper, we have deduced the upper bounds of delay and bandwidth efficient frontier based on GPS model by using network calculus theory, and established a series of upper bounds of the statistical model which suit for time delay performance of converged network. So, researching end-to-end converged network time delay performance based on network calculus has very important theoretical and practical value [4, 5].

2. Relevant Theoretical Knowledge

2.1. GPS System

GPS, Generalized Processor Sharing, is the basis of fair class scheduling algorithm and provides a guarantee to bandwidth allocation fairness. GPS is a scheduling algorithm based on continuous work. A GPS server is through the positive number parameters $\mu_1,\mu_2,\ldots,\mu_N$ said and the server is at a fixed rate for continuous work. Let $A[t,t+\tau]$ be the quality of service obtained by the flow I during a time interval $[t,t+\tau]$.The GPS model defines the flow i which is transmitted during the time interval $[t,t+\tau]$ as follows [1]:

$$\frac{A_i\left(t,t+\tau\right)}{A_j\left(t,t+\tau\right)} \geq \frac{\mu_i}{\mu_j}, j=1,2,...,N$$

The GPS algorithm has the following features:

a) Flow i can get a stable service rate and avoid being influenced by other business flows during the process of re-transmission when the average speed of the business flow i is greater than its guaranteed rate;
b) After the packet queue in buffer, the delay of the packet is defined by the maximum

Received December 21, 2013; Revised February 11, 2014; Accepted February 24, 2014

length of the business flow, because each queen is using the FCFS strategy, has nothing to do with business flow and grouping.

2.2. Network Calculus

Statistical network calculus is based on deterministic network calculus. Usually network business can tolerate a certain degree of data loss and delay[6].Through the statistical network calculus, we can make a better use of the network service so as improve the utilization ratio of resources and reflect the capabilities of the dynamic and integrated services of the network. The network calculus use cumulative function to define arrival process, service process and leave process. Ddfine $A(t)$ as arrival process, $A^{*}(t)$ as leave process and $S(t)$ as the service process of the nodes. Here is the basis of statistical network calculus which is used in this paper [7].

Definition 1. The Min-plus Convolution: The Min-plus convolution operation of functions f and g is:

$$(f \otimes g)(t) = \begin{cases} \inf\limits_{0 \le s \le t} \{ f(t-s) + g(s) \} & t > 0 \\ 0, & t \le 0 \end{cases}$$

Definition 2. Random arrival envelopes: If the cumulative flow $A[t, t+\tau]$ in any time interval $[t, t+\tau]$ satisfies the following relationship:

$$P\{A(t, t+\tau) \le \alpha(\tau)\} \ge 1 - \varepsilon$$

Call $\alpha(\tau)$ the statistical traffic envelope of the flow process, and call ε the biggest break probability.

Definition 3. Effective service curve: Given a cumulative function $A(t)$ in traffic, if the output function $D(t)$ of the traffic satisfies the following relationship:

$$P\{D(t) \ge A \otimes \beta(t)\} \ge 1 - \varepsilon$$

Say $\beta(t)$ is the effective service curve which is provided by traffic flow $A(t)$

Unlike service curve, the effective service curve is corresponding to the probabilistic lower bound of the network system service capacity.Therefore,it can describe the residual service capacity resulted from the random variations of other coexistence data flow. Whether the effective service curve can be used in statistical analysis of network service performances mainly depends on the availability of the equivalent model of the network.

Theorem 1 node series theorem: In a serial system containing n nodes, if $\beta_1(t), \beta_2(t), \cdots, \beta_N(t)$ are the arrival curve of the nodes respectively, the end-to-end effective service curve is [8]:

$$\beta = \beta_1 \otimes \beta_2 \otimes \cdots \otimes \beta_N$$

3. The Analysis of the End-to-end Delay

The total delay of GPS system based on leaky bucket model is $d = \Delta d + l$ [9], where Δd is the variable delay(also called the delay jitter).It mainly includes the delay when business flow passes through the leaky bucket regulator and the waiting delay of the business flow in the GPS system buffer. While l is the fixed delay which is mainly for the GPS system scheduling forward delay.

(1) Delay jitter

Nowadays, the business flow is usually controlled by a leaky bucket regulator. The arrival curve controlled by the leaky bucket regulator is $\alpha(t)$ and its delay jitter is $\Delta d(A,\beta)$. So we can draw that the delay jitter is equal to the waiting delay of the data flow in the buffer of the GPS system.

According to the definition of the GPS system:

$$\beta(t) = \sum_{j=1}^{N} \beta_j(t)$$

$$\frac{\beta_i(t)}{\beta_j(t)} \geq \frac{\mu_i}{\mu_j},\ j = 1, 2, \ldots, N$$

Next, let session i be used as a investigation object, we can know that:

$$\Delta d \leq \inf_{t \geq 0} \left\{ d \geq 0 : \alpha_i(t) \leq \beta_i(t+d) \right\}$$

Simultaneous above three formulas:

$$\Delta d \leq \inf_{t\geq 0} \left\{ d \geq 0 : \alpha_i(t) \leq \frac{\mu_i}{\sum_{j=1}^{N} \mu_j} \beta(t+d) \right\}$$

$$\leq \inf_{t\geq 0} \left\{ d \geq 0 : \sum_{k=1}^{c_i} \alpha_i(t) \leq \frac{\mu_i}{\sum_{j=1}^{N} \mu_j} \beta(t+d) \right\}$$

$$\leq \inf_{t\geq 0} \left\{ d \geq 0 : d \geq \frac{\frac{\sum_{j=1}^{N} \mu_j}{\mu_i} \sum_{k=1}^{c_i} r_{i,k} + b_{i,k} - Rt}{R} + T \right\}$$

And because $\frac{\sum_{j=1}^{N} \mu_j}{\mu_i} \sum_{k=1}^{c_i} r_{i,k} \leq R$,

$$\Delta d \leq T + \left(\frac{\sum_{j=1}^{N} \mu_j}{\mu_i} \sum_{k=1}^{c_i} b_{i,k} \right) / R$$

The waiting delay of session i in the buffer is defined in the above formula. That is the upper bound of delay jitter of the GPS system based on leaky bucket regulator.

(2) Delay

If l is for the fixed delay of the GPS system, the total delay will be:

$$d = \Delta d + l \leq T + l + \left(\frac{\sum_{j=1}^{N} \mu_j}{\mu_i} \sum_{k=1}^{c_i} b_{i,k} \right) / R$$

4. The Improvement on the Upper Bounds of End-to-end Statistical Delays

There are usually two ways to analyze the service performance of the data flow in the end-to-end network. The first one is that we calculate the upper bound of service performance of each node according to the corresponding relations between output and input of the adjacent nodes, and then put them together as the result of the end-to-end delay; The second one is that we use the same method to calculate the statistical upper bound of end-to-end service performance according to the tandem equivalence of several service node model [10].

In this paper, we adopt the format of the serial cascade, which can be easily extended from a single node delay boundary to the end-to-end delay boundary. In this case, the upper bound of the end-to-end delay can be concluded based on this theory:

$$d^{net} = \sum_{m=1}^{S} l_m + \sum_{m=1}^{S} T_m + \frac{S \sum_{k=1}^{c_1} b_{i,k} \sum_{j=1}^{N} \mu_j}{\mu_i \cdot \min\{r_1, r_2, \cdots, r_S\}}$$

5. Results and Analysis of the Simulation

In the end-to-end network, the number of data sources remains the same, and the scheduling queue in the cache is fixed. Comparing with the delay which has not been improved through the simulation, we can examine the effects of the corresponding parameters on the delay boundary of network.

Next, we calculate the relationships among the upper bounds d of the end-to-end statistical delay, the length of the leaky bucket b and the weight μ of the GPS model.

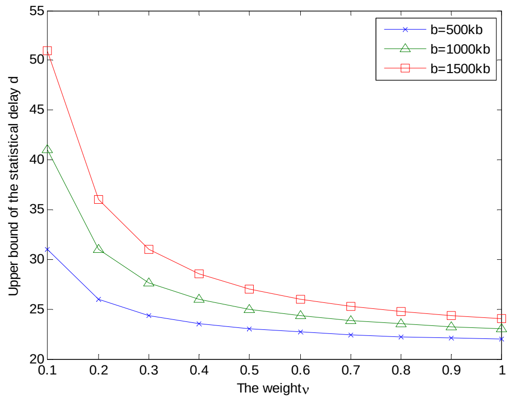

Figure 1. The Relationship of Upper Bound of the Delay, Weight and the Length of Leaky Bucket

From the Figure 1, the self-similar traffic will decrease with the increase of μ which is the GPS weight and the decrease of the length of leaky bucket.Therefore, according to the characteristics of traffic based on this feature, we can reduce the length of the leaky bucket to reduce the end-to-end delay under the premise of meet the data in the backlog. Then we can further reduce the bandwidth for the business flow to achieve the purpose of network optimization.

6. Conclusion

In this paper, we use the network calculus theory to study the behavior of GPS system based on leaky bucket regulator and deduce the statistical delay upper bound of the GPS system. Under normal circumstances, the network provides bandwidth according to the maximum rate of the data flow; this will make the network to optimization and cause waste of resources. And this article, by analyzing end-to-end delay, can provide important theoretical basis for the allocation of the network bandwidth.

Acknowledgements

This work was supported by National Natural Science Foundation of China (61271198) and Beijing Natural Science Foundation (4131003).

References

[1] Parekh AK, Gallager RG. A Generalized Processor Sharing Approach to Flow Control in Integrated Services Network: The Single Node Case. *IEEE Network*. 1993; 1(3): 344-357.
[2] Dai You, Zhge Bin, Song Huanhuan, Wang Weiming, Lan Julong. The future network research based on reconfiguration and expansible*. TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2013; 11(10): 5929-5937
[3] He Tao, chen Fucai, Qin Tao. Establishing performance evaluation model with queuing theory in NoC. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(10): 5694-5702.
[4] Boundec JYL, Thiran P. Network Calculus. Berlin: Springer Verlag. 2004.
[5] Zhang Qi-Zhi, Zhang Bin, Zhang Wei-dong. Calculation of maximum delay in switched industrial Ethernet based on network calculus*. Control and Decision*. 2005; 20(1): 117-120.
[6] Leboundec Jy, Thiran P. Network Calculus. London, Britain: Springer Verlag. 2004.
[7] Shurui Fan, Hexu Sun. Jie Li. On application method of Network Calculus to upper delay bound research. Computer Science and Information Technology (ICCSIT). 3rd IEEE International Conference on. 2010; 4: 566-570.
[8] Burchard Almut. A min-plus calculus for end-to-end statistical service guarantees. *IEEE Transactions on Information Theory*. 2006; 52(9): 4105-4114.
[9] Urvoy G, Dallery Y, Hebuterne G. CAC procedure for leaky bucket constrained sources. Performance Evaluation. 2000; 41(2): 117-132.
[10] Elwalid A, Mitra D. *Design of generalized processor sharing schedulers which statistically multiplex heterogeneous QoS classes.* Proceeding of IEEE INFOCOM'99. New York, NY, USA: IEEE. 1999: 1220-1230.