

保证业务流 QoS 的 EDF 算法

EDF Algorithm for Traffic Flows' QoS Guarantees

文 军 卢显良

(电子科技大学信息中心 成都610054)

Abstract In this paper we use token bucket traffic flows to analyze schedulable conditions under Earliest Deadline First schedulers(EDF), and present the scheduling algorithm. It can be used to ensure the bounded delay service in efficient admission control, determine whether a new flow can be admitted into network service.

Keywords QoS, Bounded delay, Earliest Deadline First schedulers(EDF), Weighted fair queuing

1 引言

网络的服务质量(QoS)对网络的吞吐量、端到端的延迟和分组的丢失率等性能参数提出了严格的要求。本文主要涉及时间延迟参数,可以用分组的端到端延迟来衡量。准许控制通过限制业务流的连接数量,能提供固定延迟的服务,但很多调度按最大资源需求分配,导致网络资源利用率低。网络延迟包含三个部分:传播延迟、传输时间和队列延迟,其中传播延迟取决于空间距离、传输介质,而传输延迟取决于数据包的大小和网络带宽,队列延迟是数据包在队列中等待的时间延迟,其中前两种的延迟是固定的。延迟主要由调度的策略决定,它是端到端延迟中变化的量,容易引起延迟发生抖动,因此,必须合理安排调度,不但要保证达到所需的延迟要求,同时还需考虑网络资源利用率。网络调度策略主要考虑下列的因素:

- 效率:能够有效地利用网络资源,支持尽可能多的连接的固定延迟服务;

- 灵活性:能够灵活地支持不同延迟的需求;

- 复杂性:算法复杂度低,使得网络不会因为算法的计算开销而使得网络等待;

- 可分析性:准许控制为保证网络 QoS,需要进行可调度性分析,并安排实现对各业务流的调度,如能保证所有的时限要求,并实现调度,则接受新连接,否则拒绝新的连接。

2 业务流特性

要对网络业务流进行调度,首先必须量化网络业务流,下面说明网络业务流的特性。假设 N 代表到达业务流的集合, $A_i(t)$ 为时刻 t 的业务流连接, $j \in N$, 则在时间 $[t_1, t_2]$ 内连接 j 的业务流的业务流表示为 $A_j[t_1, t_2]$ 。到达的数据业务流,可使用业务流约束函数 A_j^* 定义它的上界,其中 $A_j^* = 0, t < 0$; $A_j^* \geq 0, t \geq 0$ 。业务流约束函数是根据业务流的特点,使用较少的参数,描述业务流的特性,则业务流 $A_j[t, t+\tau] \leq A_j^*(\tau) \forall t \geq 0 \forall \tau \geq 0^{[2]}$ 。

为了给业务流提供 QoS 支持,需要描述每一个业务流或聚类业务流的需求。使用令牌桶系统可以精确定义通信业务流的流量特性^[1],它使用参数来表示每一个业务流,能够较容易地确定所需的资源,它包括下面的参数指标:

- 令牌桶速率 r :令牌放入令牌桶的速率,对每一个数据包,需要占用令牌桶中相当于数据包大小的令牌使用证,当一

个包的长度为 x 被服务时,则从桶中移去 x ,当数据包发送完,则占用的令牌使用证加入到空令牌桶队列中;

- 令牌桶深度大小 b :反映在短时间内通信速率超过 C 的最大数据量;

- 峰值速率 C (最大传输速率):限制包传输的最大数据传输速率。

如令牌桶大小为 b ,令牌桶速率为 r , M 为传输中的最大数据包大小,则在任意的时间间隔 T 内的速率应小于 $\min\{M + C * T, b + r * T\}$ 。

3 业务流的调度

3.1 保证固定时间延迟的方法

提供 QoS 的重要工具是调度,它主要有两种实现方法:加权公平排队和 EDF。

- 加权公平排队 根据业务流权重实施的公平排队调度,能提供细粒度的调度,假设 d_i 表示业务流 i 的最大延迟, b_i 为业务流 i 的令牌桶大小, r_i 为业务流 i 的令牌速率,业务流表示为 $A_i(t) = b_i + r_i * t$, H_i 为业务流 i 所通过的节点数, L_i 为业务流 i 路径上的最大分组长度, $L_{i,m}$ 为业务流 i 路径上链路 m 的最大分组长度, C_m 为业务流 i 通过链路的传输速率,由于能为每一业务流保证一定的带宽,则它的最大延迟上限为^[7]:

$$D_i = b_i / r_i + L_i * (H_i - 1) / r_i + \sum_{m=1}^{H_i} (L_{i,m} / C_m) \quad (1)$$

它的延迟包括了令牌桶大小引起的延迟,每个分组在每个节点上的延迟和分组传输的延迟。

- 最早时间限先服务(EDF) 一种动态的优先级调度,给每一个数据包赋予一个最晚发送的时间限,要求在此时间前完成数据包的发送,假设有 M 个业务流复用到速率为 C 的输出上,且它的时间限为 d_m ,则每一业务流 $A_m(t)$ 必须满足下面的条件:

$$\sum_{m=1}^M A_m(t - d_m) \leq C * t \quad (2)$$

$t \geq 0$, L 是链路上的 MTU,根据 EDF 调度算法,确定业务流的可调度性,才能保障业务流的延迟质量要求^[4,5]。

加权公平排队使用预留带宽,为业务流保障所需延迟的要求,虽然它算法简单,但也是比较理想的方案;在完成 EDF 中,需要在最晚时间限队列中,进行时间强度为 $O(\log Q)$ 的查找,其中 Q 是数据包队列的长度,虽然 EDF 的描述简单,但它的可调度性安排较为复杂^[3],且 EDF 被证明,如果所有的

文 军 博士研究生,研究方向为计算机网络应用和操作系统。卢显良

教授,博导,研究领域为计算机操作系统和网络,信息系统的安全性。

任务能被调度(所有任务能够满足时间限要求,且能被调度),它是最佳的调度策略,EDF 算法在保证端到端延迟的时间限制和支持的连接数量方面,其性能都远远胜过 WFQ^[8]。使用 EDF 调度策略的准许审核控制机制,能够实现高效的业务流 QoS 保障服务。

3.2 EDF 可调度性分析

假设 N 是业务流集合,业务流 i 使用 A_i^* 描述,则 EDF 调度稳定的条件是:

$$\lim_{t \rightarrow \infty} (\sum_{i \in N} A_i^*(t)) / ct < 1^{[6]} \quad (3)$$

其中 c 是恒定速率, N 是业务流集合,其中业务流 i 使用 A_i^* 描述,且它的最大延迟为 d_i , 则 EDF 能够实现调度的充要条件是:

$$ct \geq \sum_{i \in N} A_i^*(t - d_i) \quad \forall t \geq 0^{[6]} \quad (4)$$

当上面的式(3)和(4)都得到满足时,则 (A_i^*, d_i) 可被调度,它们提供了可调度性的条件,但没有提供调度性的算法,下面分析 EDF 算法。

3.3 EDF 调度算法

根据业务流的令牌桶定义,业务流 A_i^* 可表示为:

$$A_i^*(t) = \begin{cases} 0 & t < 0 \\ c_i * t & 0 \leq t < a_i \\ b_i + r_i * t & a_i \leq t \end{cases} \quad (5)$$

其中 c_i 是业务流的峰值速率, b_i 是最大突发深度, r_i 是业务流的平均速率, $a_i = b_i / (c_i - r_i)$ 是业务流以峰值速率传输的最大持续时间。

定义 $F(t) = c * t - \sum_{i \in N} A_i^*(t - d_i) \quad \forall t \geq 0$, 则式(4)条件为 $F(t) \geq 0$

由 A_i^* 的等式(5), $i \in N$, 则根据 t 的不同取值,有

$F(t) =$

$$\begin{cases} 0, & t < 0 \\ c * t - \sum_{d_i < t} C_i(t - d_i), & 0 \leq t < d_1 + a_1 \\ c * t - \sum_{i \in N} (b_i + r_i(t - d_i)) - \sum_{i > 1, d_i < t} C_i(t - d_i), & d_1 + a_1 \leq t < d_{j+1} + a_{j+1} \\ c * t - \sum_{i \in N} (b_i + r_i(t - d_i)), & d_N + a_N \leq t \end{cases} \quad (6)$$

假设一个新业务流 f , 它的参数是 (C_f, b_f, r_f, d_f) , d_f 是业务流 f 的最晚时间限, 对它进行 EDF 调度时, 它符合 EDF 的可调度性条件, 且根据 EDF 原理, 如果存在 $k \in N$, 使得 $d_{k-1} + a_{k-1} < d_f + a_f \leq d_k + a_k$, 则可把业务流 f 插入排序的业务流集合中去。此时, 当新的业务流加入集合后, 业务流的可调度性应当满足下面的条件:

$$F(d_f + a_f) \geq 0, \text{ 其中 } d_f + a_f \leq d_i \quad (7)$$

$$F(d_f + a_f) - C_f(d_f + a_f - d_f) \geq 0, \text{ 其中 } d_f < d_i + a_i \leq d_f + a_f \quad (8)$$

$$F(d_f + a_f) - C_f * a_f \geq 0, \quad (9)$$

$$F(d_f + a_f) - (b_f + r_f(d_f + a_f - d_f)) \geq 0, \text{ 其中 } d_f + a_f < d_i + a_i \quad (10)$$

通过 EDF 调度, 能够为业务流 f 分配的最小延迟是 s , 如

$s \leq d_i$, 则能够安排整个业务流集合的调度。

由条件8, 确定 $A_i^*(t - s)$, $t \in (s, s + a_f)$, 它的值小于 $C_f * a_f$, 定义 y_i , 替代 d_i , 则

$$y_i \geq (d_i + a_i) - F(d_i + a_i) / C_f, \text{ 则 } d_i \geq \max(y_1, y_2, \dots, y_i), i \in N$$

由条件10确定 $A_i^*(t - s)$, 其中 $t \in (s + a_f, \infty)$, 它的值大于 $C_f * a_f$, 定义 x_i , 替代 d_i , 则

$$x_i \geq d_i + a_i - [F(d_i + a_i) - b_f] / r_f,$$

$$\text{则 } d_i \geq \max(y_1, y_2, \dots, x_i), i \in N$$

因此有下面的算法:

```
find-minimum-delay()
/* N 是可调度业务流集合, 并且以  $(d_i + a_i)$  升序排列,  $(C_i, b_i, r_i, d_i)$ ,  $i \in N$ 
新业务流  $f(C_f, b_f, r_f, d_f)$ , 本程序实现对新业务流的安排,
 $s$  是新业务流的最小延迟时间, 它返回新业务流的最短时间 */
{
  if  $C \leq \sum_{i \in N} (r_i + r_f)$ 
    then "refuse to accept flow f and exit"
   $m_k = m_r = 0$ 
  for  $i = 1$  to  $N$  do {
    if  $F(d_i + a_i) \geq C_f * a_f$ 
      then  $x_i = d_i + a_i - (F(d_i + a_i) - b_f) / r_f$ 
      else  $y_i = d_i + a_i - F(d_i + a_i) / C_f$ 
  }
   $s = \max(s, x_i, y_i)$ 
  findout K, where  $a_k + d_k < s + a_i \leq a_{k+1} + d_{k+1}$ 
  return(s)
}
```

结论 本文使用业务流的令牌桶定义, 解决了审核控制中的 EDF 调度问题, 它能够为业务流提供所需延迟限制的服务, 达到 QoS 的要求, 同时可以使网络资源得到充分的利用。它的算法复杂度为 $O(n)$, 其中 n 为准许 EDF 调度的业务流总数。

参考文献

- Shenker S, Partridge C, Guerin R. Specification of Guaranteed Quality of Service. Internet draft-ietf-intserv-guaranteed-svc-06.txt, Aug. 1996
- Cruz R L. A Calculus for Network Delay. Part I: Network Elements in Isolation. IEEE Trans. on Information Theory, 1991, 37 (1)
- Liu C L, Layland J W. Scheduling algorithms for multiprogramming in a hard realtime environment. Journal of the Association for Computing Machinery, 1973, 20(1)
- Andrews M, Zhang L. Minimizing end-to-end delay in high-speed networks with simple coordinated schedule. In: Proc. of IEEE INFOCOM'99, New York, NY, March 1999
- Chiussi F, Sivaraman V. Achieving high utilization in guaranteed services networks using early-deadline-first scheduling. In: Proc. of the 6th IEEE/IFIP Intl. Workshop on Quality of Service, May 1998
- Liebeherr J, Wrege D, Ferrai D. Exact Admission Control for Networks with Bounded Delay Services
- Georgiadis L, Gu'erin R, Peris V, Rajan R. Efficient support or delay and rate guarantees in an Internet. In: Proc. of SIGCOMM, San Francisco, CA, Aug. 1996. 106~116
- Georgiadis L, Guerin R, Peris V, Sivarajan K N. Efficient network QoS provisioning based on per node traffic shaping. IEEE/ACM Trans. Networking, 1996, 4(4)