

面向虚拟路由器的基于历史转发开销的资源调度算法

高先明* 张晓哲 王宝生 卢泽新 马世聪
(国防科学技术大学计算机学院 长沙 410073)

摘要: 通过研究以 Xen 为代表的虚拟工具中资源调度算法, 得知它们无法保证虚拟路由器中资源共享的公平性。该文提出一种基于历史转发开销的资源调度算法, 用来确保路由器实例占用资源的公平性, 尤其物理平台的输入/输出(I/O)资源。该算法通过统计路由器实例最近的报文转发开销来计算出路由器实例调度的优先级, 确保平均吞吐量较低的路由器实例获得被调度的机会。同时, 该算法通过设置优先级计算公式中处理延迟的值, 达到对延迟敏感强的路由器实例的支持, 以及通过预测报文处理时间决定待处理报文是否丢弃, 避免不必要的系统开销。通过实验测试, 相对于 Credit 算法, 该文算法在确保虚拟路由器资源共享的公平性方面存在明显优势, 并且能为延迟敏感的路由器实例提供支持。

关键词: 网络虚拟化; 虚拟路由器; 路由器实例; 资源调度算法; 公平性

中图分类号: TP393.4

文献标识码: A

文章编号: 1009-5896(2015)03-0686-07

DOI: 10.11999/JEIT140491

Historical Forwarding Overhead Based the Resource Scheduling Algorithm for the Virtual Router

Gao Xian-ming Zhang Xiao-zhe Wang Bao-sheng Lu Ze-xin Ma Shi-cong
(School of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract: The current resource scheduling algorithms can not offer promise for the fairness of shared resources based on research on the resource scheduling algorithms in the system virtualization tools represented by Xen. This paper proposes a Historical Forwarding Overhead Based the Resource Scheduling Algorithm (HFOB_RSA) to ensure that router instances occupy determinate physical resources including I/Os that should be proportional to actual requirements. This algorithm can determine the scheduling priority of router instances by calculating the last several forwarding overhead of each router instance to make those router instances with low real throughput have an opportunity for being scheduled. Meanwhile, this algorithm also sets the value of processing delay in priority formula to provide support for those delay-sensitive router instances and determines whether or not to discard the non-disposed packets in advance by forecasting their processing time to avoid the unwanted overhead. The experimental results prove that the HFOB_RSA has superiority over Credit algorithm in terms of the fairness of shared resources in the virtual router. And HFOB_RSA also is able to provide support for the delay-sensitive router instances.

Key words: Network virtualization; Virtual router; Router instance; Resource scheduling algorithm; Fairness

1 引言

在过去的几十年里, 因特网已成为支撑社会经济发展和科技创新的重要信息基础设施, 并深刻地影响着当代人们生活与学习方式, 比如电子商务、远程教育等^[1]。当传统网络面临来自安全性、移动性和服务质量等方面的技术挑战时^[2], 通常采用“打补丁”的方式来解决所遇到的难题, 而不是采用自上而下地设计与部署全新的网络体系架构。正是这

种行为并不能有效地解决互联网自身固有的问题, 并严重阻碍了创新型网络体系架构的部署^[3]。为了破解当前互联网所面临的困境, 研究机构在借鉴系统虚拟化技术思想的基础上, 提出了将虚拟化技术应用到传统网络架构的思想, 并逐渐形成和完善了网络虚拟化技术^[4]。而路由器虚拟化技术直接关系到网络虚拟化技术的部署, 它的转发性能还影响到网络虚拟化技术的应用与推广。因此, 学术界和工业界对虚拟路由器模型以及优化机制展开深入地研究^[5]。

目前, 虚拟路由器研究方向聚焦在软件路由器方面, 即采用系统虚拟化技术实现虚拟路由器^[6,7]。如何保证虚拟路由器中资源共享的公平性

2014-04-15 收到, 2014-07-31 改回

国家 973 计划项目(2012CB315906)和高等学校博士学科点专项科研基金(20114307110006)资助课题

*通信作者: 高先明 gxm9000@163.com

是其成功部署的前提。研究人员尝试利用系统虚拟化技术中现有的资源调度算法来保证虚拟路由器资源共享的公平性，比如文献[8]中 BVT(Borrowed Virtual Time)算法^[9]、Credit 算法^[10]、SEDF(Simple Earliest Deadline First)算法^[11]以及优化算法^[12-14]。然而，这些资源调度算法设计的初衷是解决桌面系统和服务器中 CPU 公平性问题，而非输入/输出(I/O)资源的公平性。在虚拟路由器中报文转发过程主要涉及到的报文存储/读取、报文头部校验与修改以及转发表查找等，这些操作都是与 I/O 有关的。这使得虚拟路由器更多地开销花费在 I/O 资源上。另外，某些路由器实例的部署不仅要求物理资源得到保障，而且对处理延迟有着严格要求。因此，利用系统虚拟化技术中资源调度算法是无法满足虚拟路由器中资源共享公平性的需求。

因此，本文提出了一种基于历史转发开销的资源调度算法(Historical Forwarding Overhead Based the Resource Scheduling Algorithm, HFOB_RSA)来解决虚拟路由器在物理平台的资源调度方面所面临的难题。该算法通过统计每个路由器实例最近 α 次报文转发开销，计算出路由器实例调度的优先级，确定路由器实例被调度的次序，从而保证虚拟路由器资源共享的公平性。另外，该算法能预测报文处理时间，通过与路由器实例中预设的处理时延对比，预先决定报文处理方式(比如处理、丢弃等)，从而降低无效报文引入的额外转发开销。最后，通过重配置优先级计算公式中处理延迟的值，从而达到了对延迟敏感强的路由器实例的支持。通过实验证明：HFOB_RSA 能有效地保证虚拟路由器资源共享的公平性。同时，该算法能够根据路由器实例的延迟敏感度来降低报文的等待时间，为延迟敏感的路由器实例提供支持。

2 基本思想

HFOB_RSA 能依据路由器实例的报文转发历史开销数据计算出实例调度优先级，并根据优先级向路由器实例推数据包。通过控制路由器实例的调度次序来确保虚拟路由器资源共享的公平性。本节着重介绍虚拟路由器中资源调度模型以及 HFOB_RSA 的基本思想。

2.1 资源调度模型

在虚拟路由器模型中，若干个路由器实例运行在同一物理平台，并且它们相互独立、互不干扰。路由器实例作为报文处理的应用，通常位于虚拟路由器模型的顶层。它首先从下层读取报文，根据自定义报文处理方式来处理报文，最后将报文下发至下

层。为了有效地保证资源共享的公平性，虚拟路由器常常采用资源调度算法控制路由器实例读取报文的次序，是由虚拟路由器资源调度模型决定的。

在虚拟路由器资源调度模型中，报文调度模块负责将入队列中报文发送到对应的路由器实例，并将路由器实例已处理完毕的报文读取到出队列中。报文流在报文调度模块与路由器实例间的动作分为推操作和拉操作：推操作是指路由器实例被动地等待报文调度模块正在处理的数据包；拉操作是指报文调度模块主动地向路由器实例请求已处理完毕的报文，如图1所示。

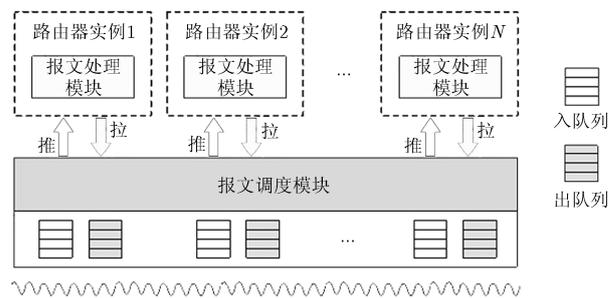


图1 虚拟路由器资源调度模型

虚拟路由器资源调度模型如何控制路由器实例调度的顺序？由于报文调度模块与路由器实例间的连接方式，路由器实例在没有接收到来自报文调度模块的数据包时，只能处于等待状态，直至报文调度模块从入队列推数据包到它。因此，报文调度模块可以通过推报文次序来控制路由器实例调度的次序，优先向优先级最高的路由器实例推数据包。

2.2 基本思想

虚拟路由器资源调度模型关键在于路由器实例优先级的计算。只有优先级最高的路由器实例才会得到报文调度模块上推的数据包，其他路由器实例则处于等待状态，直至其优先级变为最高。这使得资源调度算法中优先级的计算优劣直接影响到虚拟路由器资源共享的公平性。

优先级的计算涉及到多方面的因素：仅考虑一个因素，可能导致优先级存在偏差，无法正确反映出路由器实例对资源使用的实际状况；而某些因素可以忽略不计，考虑这些因素会增加优先级的复杂性，不能清晰地反映出优先级的主要因子。因此，合理地考虑影响因素是得出优先级计算公式的第一步。

优先级不仅要考虑路由器实例占用带宽资源的权值，而且要考虑入队列中报文等待处理时间。因为报文在入队列中等待时间越久，报文被丢弃的几

率就越大。因此, 报文等待时间越久, 对应的路由器实例优先级就应该越高, 以避免报文由于长时间等待超时而被丢弃。同时, 除了考虑以上两种因素之外, 优先级还须将路由器实例最近 α 次的报文转发开销计算在内。因为报文的长度、查表的次数、缓存(cache)的命中等都影响到路由器实例的处理开销, 如果不考虑路由器实例以往的开销, 就可能造成虚拟路由器中路由器实例占用共享资源的不公平, 无法保证资源使用率较低的路由器实例优先得到调度。此外, 优先级还需要利用处理延迟来反映路由器实例的延迟敏感度。最后, 在入队列中报文由于等待时间过程可能导致其被丢弃, 需要在报文尚未处理前通过预测报文处理时间来决定该报文是否被丢弃。

3 核心算法及特征

HFOB_RSA 中, 路由器实例调度的顺序是由它们优先级决定的。路由器实例优先级最高, 那么它就会被优先调度。本节着重探讨了 HFOB_RSA 中优先级公式、衡量算法性能的偏差量、算法的描述以及算法的可行性和时空复杂度分析。

3.1 优先级的计算公式

优先级用来表示虚拟路由器中路由器实例调度的次序。路由器实例优先级的大小直接决定它从报文调度模块获取报文的次序。同时, 优先级计算的前提条件能够有效地避免由不期望的报文处理引入的额外系统开销。优先级计算公式如式(1)所示。

$$P_i = -\lg \left(\lambda_i \frac{D_i[n]}{T_i} \frac{1}{R_{\max} \overline{D_i[\alpha]}} \right) \left(D_i[n] + f_i(\overline{D_i[\alpha]}, L) \leq T_i \right) \quad (1)$$

式中 P_i 表示第 i 路由器实例调度的优先级; λ_i 表示第 i 路由器实例带宽的权值; $D_i[n]$ 为等待因子, 表示第 i 路由器实例对应的报文等待处理的时间; T_i 表示第 i 路由器实例容忍的最大处理时延; R_{\max} 表示第 i 路由器实例的最大吞吐量; $\overline{D_i[\alpha]}$ 为历史因子, 表示第 i 路由器实例最近 α 条报文转发开销的均值; $f_i(\overline{D_i[\alpha]}, L)$ 为预测处理时间, 表示第 i 路由器实例下一报文预测处理时间, L 表示报文的长度。

从式(1)可以看出: 当路由器实例对应的报文等待处理时间与预测处理时间之和不超过该路由器实例容忍的最大处理时延时, 优先级公式才会被调用。路由器实例调度的优先级不仅与它分配带宽的权值和报文的等待处理时间呈正比, 而且与它设置的最大容忍延迟、最大吞吐量以及它最近 α 条报文转发开销的均值呈反比。其中, 优先级与带宽 λ_i 呈

正相关, 带宽的权值越大, 那么路由器实例的调度优先级就越高。 $D_i[n]$ 是表示入队列中报文等待时间, $D_i[n]$ 越大, 对应的路由器实例就越容易被调度。同时, 入队列中报文的等待时间 $D_i[n]$ 加上预测处理时间超过 T_i 门限时就会被丢弃, 从而避免不期望的处理开销。 T_i 值表示路由器实例容忍的最大处理延迟, T_i 越大, 路由器实例对延迟敏感度越低, 调度优先级也就越小。 R_{\max} 反映了路由器实例处理报文的能力: 处理报文能力越强, 它支持的最大吞吐量就越高。考虑到路由器实例间调度公平性, $\overline{D_i[\alpha]}$ 反映路由器实例最近的转发开销, 与优先级呈反比, $\overline{D_i[\alpha]}$ 越大, 那么其优先级就越低, 从而确保平均吞吐量小的路由器实例也有机会获得调度。

(1)历史因子 历史因子 $\overline{D_i[\alpha]}$ 是指路由器实例最近 α 次报文转发开销计算出均值。历史因子计算公式为

$$\overline{D_i[\alpha]} = \frac{\sum_{i=1}^{\alpha} t_i[n]}{\alpha} \quad (2)$$

其中, $t_i[n]$ 表示第 i 路由器实例最近第 n 次转发开销。 α 值表示路由器最近报文转发开销的次数, 是历史因子的阈值。 α 值直接影响到历史因子的表现: 阈值 α 过大, 历史因子 $\overline{D_i[\alpha]}$ 无法准确地反映出路由器实例当前时间段的平均吞吐量, 尤其报文流波动明显时; 如果阈值 α 过小, 历史因子 $\overline{D_i[\alpha]}$ 仅能反映最近时间段内路由器实例状况, 无法有效地保证资源共享的公平性。

(2)前提条件 优先级计算的前提条件能在报文尚未处理前决定报文行为(包括处理或丢弃)。具体来讲, 通过判断预测报文处理时间和报文等待时间之和是否超过路由器实例预设处理时延: 如果没有超过, 则计算路由器实例优先级; 否则直接丢弃该报文, 避免该报文引入的不必要的开销。前提条件 $D_i[n] + f_i(\overline{D_i[\alpha]}, L) \leq T_i$ 中, $D_i[n]$ 表示报文的等待处理时间, T_i 表示路由器实例预设的处理时延, 是已知的。前提条件判断待处理报文是否超时取决于 $f_i(\overline{D_i[\alpha]}, L)$, 与历史因子 $\overline{D_i[\alpha]}$ 和报文长度 L 两个参数相关。 $f_i(\overline{D_i[\alpha]}, L)$ 表示预测的报文处理时间, 其计算公式为

$$f_i(\overline{D_i[\alpha]}, L) \approx k \cdot L \cdot \frac{\overline{D_i[\alpha]}}{L_i[\alpha]} \quad (3)$$

式中 k 为相关系数, 与物理平台有关; $\overline{L_i[\alpha]}$ 表示最近 α 个报文的平均长度, L 表示当前待处理的报文长度。从式(3)可以看出: $f_i(\overline{D_i[\alpha]}, L)$ 与报文的长度呈正相关, 报文长度越长, 预测报文处理时间就越长。利用该公式能预测出报文处理时间。

3.2 偏差量

偏差量是指虚拟路由器中路由器实例实际占用的共享资源与理论占用的共享资源的均方差，用于衡量资源调度算法的公平性，偏差量计算公式为

$$\vartheta = \sum_{i=1}^N \left(\frac{T_{\text{real}}(i) - T_{\text{theory}}(i)}{T_{\text{theory}}(i)} \right)^2 \quad (4)$$

式中 ϑ 表示资源调度算法的偏差量； $T_{\text{real}}(i)$ 表示路由器实例 i 实际的处理报文开销； $T_{\text{theory}}(i)$ 表示路由器实例 i 理论的处理时间。从式(4)可以看出：偏差量 ϑ 的大小取决于路由器实例占用共享资源的公平性。如果偏差量 $\vartheta \rightarrow 0$ ，表示该资源算法能够完全保证共享资源的公平性。如果偏差量 ϑ 越大，表示该资源调度算法在保证共享资源公平性方面表现就越差。因此，可以利用偏差量 ϑ 来评估阈值 α 对 HFOB_RSA 性能的影响，从而决定阈值 α 的取值范围。当偏差量超过某个设定值，可以通过增加阈值 α 来保证虚拟路由器中资源公平性。

3.3 算法具体描述

HFOB_RSA 中历史因子反映当前路由器实例的平均处理能力，并尽可能地保证各个路由器实例占用资源的公平性。相对于其他资源调度算法，它能根据路由器实例历史转发开销，确保那些平均吞吐量低的路由器实例更易获取被调度的机会，从而保证虚拟路由器资源共享的公平性。HFOB_RSA 通过优先级决定路由器实例调度的次序，并能预判待处理报文处理方式。算法的具体描述如表 1 所示。

表 1 算法的具体描述

```

HFOB_RSA 算法描述:
Program HFOB_RSA(输入: 入队列不为空
                  输出: 处理完报文)
  DoWhile(遍历入队列)
Next: 计算  $D_i[n] + f_i(\overline{D}_i[\alpha], L)$ ;
      IF  $D_i[n] + f_i(\overline{D}_i[\alpha], L) \leq T_i$  THEN
          计算  $\overline{D}_i[\alpha]$ ;
          计算  $P_i$ ;
      ELSE
          丢弃报文, 并读取下一报文;
      IF 队列不为空 THEN
          GOTO Next;
      END IF
  END IF
END While
计算  $\text{Max}\{p_0, p_1, \dots, p_n\}$ ;
将报文推到优先级最高的路由器实例对应的入队列;
END HFOB_RSA

```

通过 HFOB_RSA 描述可知：只有入队列存在报文时才触发该调度算法，否则，报文调度模块一直处于空转状态。在该资源调度算法中，首先利用前提条件预测待处理报文是否超时，如果存在超时情况，则读取入队列中下一条报文；否则计算出路由器实例的历史因子和调度优先级。最后，比较出优先级最高的路由器实例，并将报文上推到该路由器实例中。HFOB_RSA 与其他调度算法最大差别在于前者通过统计路由器实例的历史转发开销来更加精确地确定路由器实例调度次序，从而确保虚拟路由器资源共享的公平性。

4 实验与分析

本小节主要利用实验来测量 HFOB_RSA 和 Credit 算法的性能，验证 HFOB_RSA 在保证资源公平性方面和支持延迟敏感服务方面具有优势。

4.1 实验环境

本实验拓扑包括一台客户端和一台部署虚拟路由器的服务器，客户端采用 lenovo ThinkCentre，具有双核的 2.70 GHz 奔腾处理器，4 G 内存，和 2 个 1 Gbps 接口，虚拟路由器部署在 lenovo Think Centre，具有双核的 2.13 GHz Xeon 处理器，4 G 内存，和 2 个 1 Gbps 接口，如图 2 所示。客户端采用 Lan Traffic^[15]发送报文长度不同的用户数据报协议(User Datagram Protocol, UDP)报文，比如 64 byte、128 byte 等长度。同时，客户端建立 UDP 服务器，用于监听虚拟路由器转发的 UDP 报文，并记录数据包的相关信息(包括报文 ID 号、报文长度、报文发送时间戳和报文接收时间戳)。虚拟路由器采用类似 Click 软件路由器实现。在虚拟路由器中，每个路由器实例对应于一个类似 Click 软件路由器实例。在本实验中，类似 Click 软件路由器实例处理 IPv4 协议报文，实现报文从一个端口接收和另一端口发送。

本文实验有 3 个目标：(1)探讨 α 值对 HFOB_RSA 性能的影响；(2)探讨报文长度对 HFOB_RSA 与 Credit 算法的影响；(3)验证 HFOB_RSA 支持延迟敏感的路由器实例。

4.2 前提条件中 k 值的测量

前提条件中 k 值是与物理平台的处理能力和报

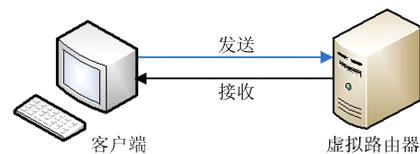


图2 实验环境中网络拓扑结构

文长度有关的,是未知量。为了求出该变量的取值范围,本实验测试了虚拟路由器在该平台上处理不同长度报文(包括 64 byte, 128 byte, 256 byte, 512 byte, 768 byte, 1024 byte, 1500 byte)时,单条报文的处理时间,如图 3 所示。

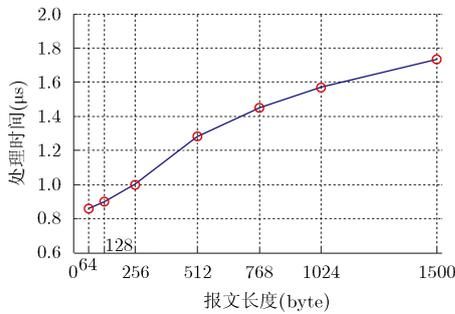


图3 不同报文长度时的处理时间

从图 3 可以看出,随着报文长度的增加,其报文的处理时间也相应地增加。报文的处理时间与它的长度存在一定的相关性。由式(3)的定义可知: $f_i(\overline{D}_i[\alpha], L)$ 可以表示下一报文的处理时间,那么进一步推导出 k 值的计算公式为

$$k = \left(\frac{T_i[\text{len}]}{T_i[\text{Blen}]} - 1 \right) / (L_i[\text{len}] - L_i[\text{Blen}]) \quad (5)$$

$T_i[\text{len}]$ 表示该长度报文的处理时间; $T_i[\text{Blen}]$ 表示基长度报文的处理时间,本实验是以 64 byte 报文作为参照值。利用式(5)和图 3 的实验数据,计算出 k 的取值范围为区间 $[0.000716, 0.001106]$ 。如果 k 值取值过小,会使报文的预测处理时间小于实际处理时间,可能导致部分不满足前提条件的报文而实际被路由器实例处理,从而引入额外的系统开销。因此,实验中 k 值设定为该区间的上限,即 0.001106。

4.3 α 对 HFOB_RSA 性能的影响

虚拟路由器部署 3 台路由器实例:路由器实例 1、路由器实例 2 和路由器实例 3。它们的处理能力分别为转发 64 byte 报文的速率是 150 Mbps, 150 Mbps 和 300 Mbps。同时,配置 3 台路由器实例的报文长度比值为 1:2:2,带宽比值为 1:1:2,并设置 HFOB_RSA 中 α 值为 1, 10, 20, 30, 40, 50。当整个虚拟路由器处理报文数目为 1000 时,通过利用式(5)来计算出 HFOB_RSA 的偏差量,以及统计出 3 个路由器实例处理的报文数目,如表 2 所示。

从表 2 中看出,在 6 种不同场景,路由器实例的处理报文数目存在差异。当偏差量值小于 0.02795 时,路由器实例处理的报文数目不再发生改变。随着偏差量逐渐减少,路由器实例处理报文

表2 HFOB_RSA的偏差量及路由器实例处理的报文数

α	偏差量	路由器实例 1	路由器实例 2	路由器实例 3
1	0.06458	320	160	520
10	0.03435	300	180	520
20	0.02823	320	220	460
30	0.02795	320	240	440
40	0.02795	320	240	440
50	0.02795	320	240	440

数目趋于恒定,这说明 HFOB_RSA 在保证资源公平性方面表现也越来越好。实验中,偏差量为 0.02795 时, HFOB_RSA 算法在保证资源公平性方面已经满足于本文的需求(在本实验中 $\alpha = 30$)。

4.4 报文长度的差异

在虚拟路由器中部署了两个路由器实例:路由器实例 A 和路由器实例 B。它们两者的处理能力相同,为转发 128 byte 报文的速率为 300 Mbps。同时,配置路由器实例 A 转发的报文长度分别为 64 byte、128 byte 和 256 byte,路由器实例 B 的转发报文长度为 128 byte。通过测量两台路由器实例在实际运行过程中占用的共享资源来探讨报文长度对 Credit 算法和 HFOB_RSA 性能的影响,如图 4 所示(横轴表示时间,纵轴表示路由器实例 A 与路由器实例 B 占用共享资源的比值)。

通过分析图 4 可以发现:当路由器实例 A 转发的报文长度为 64 byte 时, Credit 算法中路由器实例 A 占用的共享资源几乎是路由器实例 B 的两倍, HFOB_RSA 使路由器实例 A 与路由器实例 B 占用共享资源的比值在 1.2 左右;在路由器实例转发 128 byte 报文时,两种算法的表现基本一致;当路由器实例 A 转发 256 byte 报文时, Credit 算法中路由器实例 A 占用的资源仅为路由器实例 B 的大约 65%,而 HFOB_RSA 则能将路由器实例 A 与路由器实例 B 占用共享资源的比值提高到 80%。

4.5 延迟敏感度的差异

虚拟路由器部署两台路由器实例:路由器实例 C 和路由器实例 D。它们的处理能力相同,转发的报文长度相同,为转发 64 byte 报文的速率是 300 Mbps。同时,配置路由器实例 C 的延迟敏感度为 1,路由器实例 D 的延迟敏感度为 2(即式(1)中 T_i 值为 1:2),使得路由器实例 D 的延迟敏感程度要强于路由器实例 C。统计两台路由器实例中报文等待时间的均值,如图 5 所示(横轴表示时间,纵轴表示报文等待时间)

从图 5 中可以清晰地看出:由于路由器实例 C

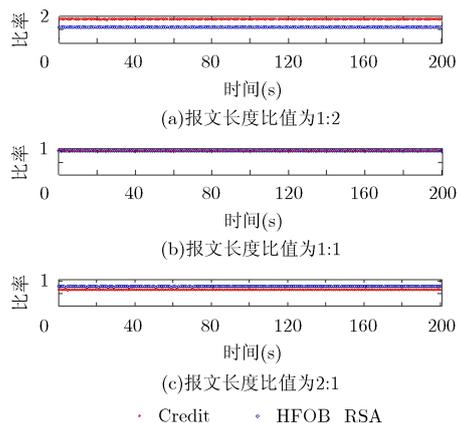


图4 两台路由器实例的占用资源比值

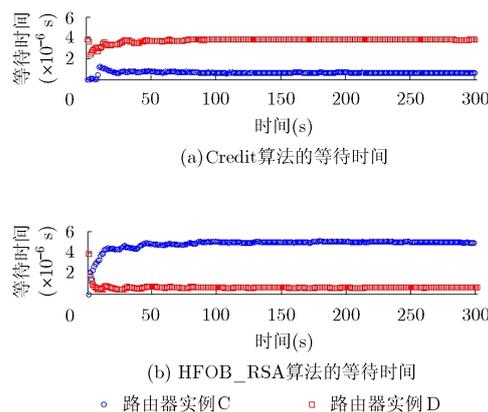


图5 路由器实例中报文等待时间

的报文优先于路由器实例 D 的报文到达，Credit 算法依据先来先处理的原则，优先处理路由器实例 C 的报文，所以导致路由器实例 D 中报文等待时间是路由器 C 的 4 倍。而 HFOB_RSA 能够根据延迟敏感的需求，优先调度延迟敏感度高的路由器实例 D，这样使得路由器实例 D 中报文等待时间明显低于路由器实例 C 中报文等待时间。

5 结束语

由于虚拟路由器支持多个路由器实例并行运行在同一物理平台，这使得多个路由器实例同时访问物理资源时，就会产生资源竞争的问题。如何保证虚拟路由器中资源共享的公平性是当前迫切需要解决的问题。本文针对现有的资源调度算法存在的不足，提出了一种 HFOB_RSA。该算法能利用路由器实例最近 α 次的转发开销来确定它们的调度次序，从而确保共享资源的公平性。通过实验分析，相比较 Credit 算法，HFOB_RSA 在保证资源公平性方面具有明显的优势。同时，HFOB_RSA 可以为延迟敏感高的路由器实例提供支持。

然而，HFOB_RSA 也存在一些问题：(1) α 值取值多少才能满足需求，过大就会引入较多的系统开销，从而降低虚拟路由器的整体性能，较少则无法保证资源的公平性；(2) 从图 4 可以看出 HFOB_RSA 并不能严格地保证资源的公平性，只是相对其他资源调度算法而言，能较好地保证资源公平性。

参考文献

- [1] Anderson T, Peterson L, Shenker S, *et al.* Overcoming the Internet impasse through virtualization[J]. *Computer*, 2005, 38(4): 34-41.
- [2] Clark D, Braden R, Sollins K, *et al.* New arch: future generation internet architecture[R]. Massachusetts Institute of Technology Cambridge Lab for Computer Science, Massachusetts, USA, 2004.
- [3] Turner J S and Taylor D E. Diversifying the internet[C]. Proceedings of the Global Telecommunications Conference, Saint Louis, USA, 2005, 2: 755-760.
- [4] Guimaraes P H V, Ferraz L H G, Torres J V, *et al.* Experimenting content-centric networks in the future internet testbed environment[C]. Proceedings of the 2013 IEEE International Conference on Communications Workshops, Budapest, Hungary, 2013: 1383-1387.
- [5] Wang A, Iyer M, Dutta R, *et al.* Network virtualization: technologies, perspectives, and frontiers[J]. *Journal of Lightwave Technology*, 2013, 31(4): 523-537.
- [6] Fischer A, Berl A, and de Meer H. Virtualized networks based on system virtualization[C]. Proceedings of the 2nd GI/ITG KuVS Workshop on The Future Internet, Munich, Germany, 2008: 44-47.
- [7] Ahm C, Lee S H, and Yoo C. Performance analysing for virtual router on xen hypervisor[C]. Proceedings of the 8th International Symposium on Embedded Technology, Daegu, Korea, 2013: 1-2.
- [8] Bourguiba M, Haddadou K, and Pujolle G. Evaluating and enhancing xen-based virtual routers to support real-time applications[C]. Proceedings of the Consumer Communications and Networking Conference (CCNC), Nevada, USA, 2010: 1-5.
- [9] Duda K J and Cheriton D R. Borrowed-virtual-time (BVT) scheduling: supporting latency-sensitive threads in a general-purpose scheduler[C]. Proceedings of the ACM SIGOPS Operating Systems Review, New York, USA, 1999: 261-276.
- [10] Kim N, Cho J, and Seo E. Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems[J]. *Future Generation Computer Systems*, 2014, 32: 128-137.
- [11] Gu Z and Zhao Q. A state-of-the-art survey on real-time issues in embedded systems virtualization[J]. *Journal of Software Engineering & Applications*, 2012, 5(4): 277-290.

- [12] Gulati A, Merchant A, and Varman P J. mClock: handling throughput variability for hypervisor IO scheduling[C]. Proceedings of the 9th USENIX conference on Operating systems design and implementation, British Columbia, Canada, 2010: 1-7.
- [13] Liao G, Guo D, Bhuyan L, *et al.* Software techniques to improve virtualized I/O performance on multi-core systems [C]. Proceedings of the 4th ACM/IEEE Symposium on Architectures for Networking and Communications Systems, San Jose, USA, 2008: 161-170.
- [14] Ongaro D, Cox A L, and Rixner S. Scheduling I/O in virtual machine monitors[C]. Proceedings of the fourth ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments, Seattle, USA, 2008: 1-10.
- [15] Packet Data Systems Ltd. Lan Traffic[OL]. http://www.pds-test.co.uk/products/ip_generation.html, 2014.4.
- 高先明: 男, 1988 年生, 博士, 研究方向为新型网络体系结构、网络虚拟化技术.
- 张晓哲: 男, 1976 年生, 硕士生导师, 副研究员, 研究方向为新型网络体系结构、路由器体系架构.
- 王宝生: 男, 1970 年生, 博士生导师, 研究员, 研究方向为计算机网络与通信、网络安全.
- 卢泽新: 男, 1964 年生, 硕士生导师, 研究员, 研究方向为计算机网络与通信、网络安全.
- 马世聪: 男, 1986 年生, 博士, 研究方向为新型网络体系结构、网络虚拟化技术.