

基于网络演算的聚合模型分析方法及其评估

龙彦辰¹, 沈海斌¹, 鲁中海²

(1. 浙江大学超大规模集成电路设计研究所, 浙江杭州 310027; 2. 瑞典皇家工学院, 斯德哥尔摩 16440)

摘 要: 在资源共享的通信媒介中, 例如总线、交叉矩阵、网络, 聚合模型的应用十分普遍. 经由聚合节点发送的数据包的延迟上界可以通过网络演算理论得到, 然而对于这种延迟上界紧致性的研究一直以来都是开放性问题. 本文基于网络演算理论, 讨论了单节点、多节点聚合模型的分析方法, 包括两种到达曲线模型、两种等价服务曲线分析方法. 同时以实验手段对延迟上界的紧致性进行了经验性的评估. 最后作为结论, 本文归纳了各种分析模型的质量, 总结了各个参数是如何影响延迟上界的紧致性的.

关键词: 网络演算; 聚合模型; 延迟上界; 紧致性

中图分类号: TP302 **文献标识码:** A **文章编号:** 0372-2112 (2018)08-1815-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2018.08.003

Analysis and Evaluation of Delay Bounds for Multiplexing Models Based on Network Calculus

LONG Yan-chen¹, SHEN Hai-bin¹, LU Zhong-hai²

(1. Institute of VLSI Design, Zhejiang University, Hangzhou, Zhejiang 310027, China;

2. KTH Royal Institute of Technology, Stockholm 16440, Sweden)

Abstract: In resource-sharing communication media such as buses, crossbars and networks, multiplexings are inevitable. While sending packets over a multiplexing node, the worst-case delay bound can be computed using network calculus. The tightness of such delay bound remains an open problem. This paper studies different analysis approaches for multiplexing models, from the single multiplexing node to multi-flow-multi-node model, applying two traffic arrival models, and two service properties when getting equivalent service curves. We analyze per-flow delay bounds with different models, then empirically evaluate the tightness of the delay bounds. Our results show the quality of different analysis models, and how influential each parameter is to tightness.

Key words: network calculus; multiplexing model; delay bound; tightness

1 引言

在现代片上多核系统当中, 片上互联起到了越来越关键的作用^[1,2]. 如果将整个系统划分为通信、计算、存储三大子系统, 为了满足实时性的需求, 保证通信子系统的性能显得至关重要^[3]. 为了准确预测系统性能, 为资源共享的情景建立适当的分析模型, 从而获得性能上界的分析必不可少.

通道复用是一种常见的资源共享方式. 例如片上网络(Network-on-Chip, NoC)中, 数据包通过多跳的方式在路由器之间的通道中传输, 从源节点发送到目标节点. 我们将这种单播数据流简称作流. 一个流的数据包很可能与其他流的数据包发生竞争, 从而导致路径阻塞. 因此我们要给最大数据包延迟提供上界以保证通信性能. 然而

前人在这方面的研究还远远不够, 尤其在对复用分析模型的紧致性问题上并没有进行充分的评估.

本文在网络演算(Network Calculus)理论的基础上分析了聚合模型的最坏情况延迟上界问题. 我们分别讨论了两种不同的到达模型, (b, r)模型和 TSPEC 模型; 以及两种不同的聚合节点服务曲线分析法, 分别为剩余服务量分析法(Left-over service Property, LP)和带宽分配分析法(Isolation Property, IP). 按照网络的复杂度分类, 我们首先讨论了单流-多节点延迟上界的求取方法, 后讨论了多流-单节点的聚合模型, 进而得出多流-多节点聚合模型的端到端延迟上界的通用分析方法. 通过研究, 我们分析归纳了不同情况下各模型所具有的优势, 并总结出了一般情况下紧致性上界出现的注入信息流和服务节点能力的配置. 这有效地帮助我

们缩短了软件模拟时长,预测何时、何种配置下,可能有最坏延迟情况的出现.

2 相关工作

网络演算是一种用于性能保证分析的理论,主要基于对流量的到达曲线和服务器的服务曲线的高度抽象建模.网络演算最初应用于通信网络领域,经过大批学者不懈的研究,该理论如今被广泛应用于嵌入式系统,片上系统及片上网络之中.

Cruz 首先在他的文章^[4]中提出了这种以积分的方法求得延迟上界的理论,以及基于注入流的突发传输量和平均注入速率的到达曲线模型.随后服务曲线的定义也被提出^[5].Boudec 和 Thiran 系统地总结了前人的工作,并给出了网络演算理论与极大极小代数的关系^[6].

在解决资源共享建模的难题上,文献^[6]给出了各种不同的聚合节点的延迟上界.文献^[7]中,作者将网络演算理论应用于片上网络,对三种基本干扰模型做了等价服务曲线分析,并计算出单个流的延迟上界.

另外,平均性能估算也是片上网络的重要研究议题.基于泊松流量到达假设和马尔可夫过程,研究者们给出了不同的平均延迟估算模型^[8-10].

3 网络演算基础

表 1 列出了本文中使用的数学符号、表达式的定义.

表 1 数学符号与表达式定义

数学符号	定义
f_i	第 i 路信息流
R_i	第 i 路信息流的累积到达函数
α_i	第 i 路信息流的输入到达曲线
$\gamma_{r,b}$	突发传输量为 b 且平均速率为 r 的到达曲线, $\gamma_{r,b} = rt + b, t$ 表示事时间
β^{N_j}	节点 N_j 的服务曲线
$\beta_i^{N_j}$	信息流 f_i 在节点 N_j 的等价服务曲线
$\hat{\beta}_{sys}$	目标流的端到端等价服务曲线
$\beta_{R,T}$	最低速率为 R 且最大延迟为 T 的延迟-速率服务曲线 $\beta_{R,T} = R[t - T]^+ = \begin{cases} R(t - T), & t > 0 \\ 0, & \text{otherwise} \end{cases}, t$ 表示时间
\bar{D}	目标流的端到端延迟上界.其中 \bar{D}_{br} 表示 (b, r) 到达曲线模型计算所得延迟上界, \bar{D}_{isp} 表示 TSPEC 到达曲线模型计算所得延迟上界
$h(\alpha, \beta)$	曲线 α 和 β 的最大水平距离
ϕ_i	信息流 f_i 获得的权重值
\otimes	最小加代数域卷积: $f \otimes g(t) = \inf_{0 \leq s \leq t} \{f(t-s) + g(s)\}$

3.1 到达曲线 (b, r) 模型和 TSPEC 模型

根据文献^[6]对到达曲线的定义,当且仅当对于任意 $\forall s < t, R(t) - R(s) \leq \alpha(t - s)$ 都成立时, α 为流 R 的一个到达曲线.以仿射函数

$$\gamma_{r,b} = \alpha(t) = \begin{cases} rt + b, & \text{if } t > 0 \\ 0, & \text{otherwise} \end{cases}$$

为到达曲线的信息源,就是到达曲线的 (b, r) 模型,其中 b 为最大突发传输量, r 为平均速率.在实际分析中,有时还需考虑传输过程中的峰值行为.因而文献^[11]提出了 TSPEC (Traffic Specification) 模型.该模型由四个参数 (p, M, r, b) 组成,其中 M 为数据包的最大长度, p 为峰值传输速率,最大突发传输量 b 和平均速率 r 与 (b, r) 模型对应.到达曲线的数学表达式为

$$\alpha(t) = \min(M + pt, rt + b)$$

如图 1 中的红色实线部分.

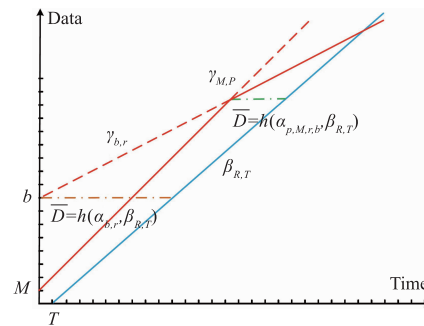


图 1 (b, r) 与 TSPEC 的比较

3.2 服务曲线:延迟-速率服务节点

假设流 f 在系统的输入端和输出端分别由累积函数 R 和 R^* 表示,当且仅当:(1) β 是一个广义增函数;(2) $\beta(0) = 0$;(3) $R^* \geq R \otimes \beta$ 时,系统向这路数据流提供一个服务曲线 β .若

$$\beta = R[t - T]^+ = \begin{cases} R(t - T), & \text{if } t > T \\ 0, & \text{otherwise} \end{cases}$$

那么 β 被称为是最小服务速率为 R 、最大延迟为 T 的延迟-速率服务曲线,由 $\beta_{R,T}$ 表示^[6].

3.3 延迟上界

假设流 f 的到达曲线为 α ,系统的服务曲线为 β ,那么 f 的延迟上界 \bar{D} 为曲线 α 和 β 的最大水平距离 $h(\alpha, \beta)$,如图 1 中的点画线.若 α 为仿射函数 $\gamma_{r,b}$, β 为延迟-速率服务 $\beta_{R,T}$,流 f 的延迟上界为^[6]:

$$\bar{D}_{br} = T + \frac{b}{R}, r \leq R$$

若 α 用 TSPEC 模型 (p, M, r, b) 描述,服务曲线不变, f 的延迟上界为^[6]:

$$\bar{D}_{isp} = \frac{M + \frac{b - M}{p - r}(p - R)^+}{R} + T, r \leq R$$

如图 1,通常 \bar{D}_{br} 的值大于等于 \bar{D}_{isp} 的值. 当 $p > r$ 时, R 与 r 的差值越大, \bar{D}_{isp} 和 \bar{D}_{br} 的差值也越大.

4 聚合模型及其分析方法

4.1 单流模型

对于信息流经过多跳才能到达目的地的情况,我们使用等价服务曲线的方法分析端到端延迟上界^[7]. 首先,将流在途经的各节点的服务曲线进行极小代数域上的卷积计算,得到端到端等价服务曲线 $\hat{\beta}_{sys}$, 然后通过该流的注入到达曲线,计算出端到端延迟上界 \bar{D} . 如图 2 所示,单流-多节点分析过程如下:

$$\hat{\beta}_{1_sys} = \beta^{N_1} \otimes \beta^{N_2} \cdots \otimes \beta^{N_m}$$

$$\bar{D}_1 = h(\alpha_1, \hat{\beta}_{1_sys})$$

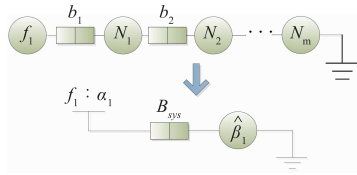


图2 单流-多节点模型

4.2 多流-单节点模型

4.2.1 聚合节点的分类

文献[7]中,片上网络的聚合模型被总结为链路共享,缓存共享和流控回路共享三类. 无论哪种共享方式,在分析时,我们首先将这个多流聚合的节点转化为多个独立的“单流-(单)等价服务节点”的模型,再分别求取它们的等价服务曲线,如图 3 所示. 我们需要求取延迟上界的流定义为目标流,其它则为干扰流.

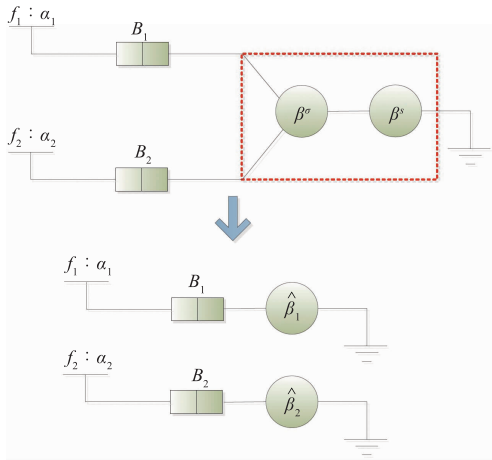


图3 单流-多节点模型

4.2.2 剩余服务量分析方法

剩余服务量分析法的核心思想在于目标流在单个节点所能得到的服务不仅与节点本身的服务能力相关,也依赖于干扰流的负载量.

假定目标流 f_1 和干扰流 f_2 分别服从到达曲线 α_1 , α_2 , 由任意的仲裁方式通过服务节点 N_1 并成为聚合流, N_2 服务曲线为 $\beta_{R,T}$. 根据文献[6]可证,无论 α_2 是由 (b,r) 模型 γ_{r_2,b_2} 描述,亦或由 TSPEC 模型 (p,M,r_2,b_2) 描述, $\hat{\beta}_{sys_lp}(t) = [\beta_{R,T}(t) - \alpha_2(t)]^+$ 总是一个广义增函数. 因此, $\hat{\beta}_{sys_lp}(t)$ 是 f_1 的一个服务曲线. 在实际中,总有 $R \leq p = 1$, 所以 f_1 在节点 N_1 处的等价服务曲线为:

$$\hat{\beta}_{sys_lp}(t) = (R_{sys_lp})[t - T_{sys_lp}]^+$$

其中,

$$\begin{cases} T_{sys_lp} = \frac{b_2 + RT}{R - r_2} \\ R_{sys_lp} = R - r_2 \end{cases}$$

由此,分别根据 α_1 的 (b,r) 模型和 TSPEC 模型,得到 f_1 在节点 N_1 处的延迟上界 \bar{D}_{br_lp} 和 \bar{D}_{isp_lp} :

$$\bar{D}_{br_lp} = T_{sys_lp} + \frac{b_1}{R_{sys_lp}}, r_1 \leq R_{sys_lp}$$

$$\bar{D}_{isp_lp} = \frac{M + \frac{b_1 - M}{p - r_1}(p - R_{sys_lp})^+}{R_{sys_lp}} + T_{sys_lp}, r_1 \leq R_{sys_lp}$$

4.2.3 带宽分配分析方法

在一些时分多路复用的仲裁方式中,每一路等候队列会被分配一个相应的时间段进行传输. 例如片上常用的加权轮询 (Weighted Round Robin, WRR), 每一路流被分配的时间段长度 (以时钟周期数计) 等于该路等候队列的权重 ϕ_i . 在每次轮询中,仲裁器允许流 f_i 在它的时间段内传输至少 ϕ_i 个数据单位. 如果 f_i 到达输入端口时刚刚错过它的时间段,那么 f_i 就会遇到它在该节点最长的等待时间.

假定该节点的服务速率为 R_σ , 则其对 f_i 提供的服务保障为 $R_\sigma \frac{\phi_i}{\sum_{j=1}^M \phi_j}$, M 为此节点等候队列的数量. f_i 在此节点可遇的最大延迟为 $\frac{\sum_{j \neq i} \phi_j}{R_\sigma}$, 用突发延迟函数表示为^[7]:

$$\delta_{\frac{\phi_i}{R_\sigma}}(t) = \begin{cases} +\infty, & \text{if } t > \frac{\sum_{j \neq i} \phi_j}{R_\sigma} \\ 0, & \text{otherwise} \end{cases}$$

在这种分析方法中,将节点对各路流的服务分别当作多个独立的、不受其他干扰流影响的等价服务来看待;而每个等价服务曲线的计算,直接与各路流所被保证的带宽密切相关. 因此,我们称这种分析方法为带宽分配法 (Isolation Property, IP). 以两路流聚合为例,目标流 f_1 的等价服务曲线为:

$$\hat{\beta}_{sys_ip} = \frac{\phi_1}{\phi_1 + \phi_2} \beta^\sigma \otimes \beta \otimes \delta_{\frac{\phi_2}{R_\sigma}}$$

化简后为一个延迟-速率服务曲线:

$$\hat{\beta}_{\text{sys_ip}}(t) = R_{\text{sys_ip}} [t - T_{\text{sys_ip}}]^+$$

其中,

$$\begin{cases} T_{\text{sys_ip}} = T_{\sigma} + T + \frac{\phi_2}{R_{\sigma}} \\ R_{\text{sys_ip}} = \frac{\phi_1}{\phi_1 + \phi_2} (\min\{R_{\sigma}, R\}) \end{cases}$$

由此,推导出带宽分配法所得的单节点延迟上界. 对 (b, r) 模型:

$$\bar{D}_{br_ip} = T_{\text{sys_ip}} + \frac{b_1}{R_{\text{sys_ip}}}, r_1 \leq R_{\text{sys_ip}}$$

对 TSPEC 模型:

$$\bar{D}_{isp_ip} = \frac{M + \frac{b_1 - M}{p - r_1} (p - R_{\text{sys_ip}})^+}{R_{\text{sys_ip}}} + T_{\text{sys_ip}}, r_1 \leq R_{\text{sys_ip}}$$

4.2.4 两种分析方法的应用情况

在以上两种分析方法中,剩余服务量分析法的推导过程假定了任意的仲裁机制,而带宽分配法则基于时分多路复用的仲裁机制及其参数,因此,它们的适用范围也有所差别.

剩余服务量分析法为一种普遍适用的分析方法,即便我们缺少多路流的聚合模型、仲裁方式的信息,依然可以使用此方法. 它的意义在于,在现阶段的研究中,一些聚合情况只能使用此方法进行分析,如图 4 中节点 N_2 处缓存共享的聚合流,流 f_1 的等价服务曲线 $\hat{\beta}_1^{N_2}$ 只能通过 f_2 在 N_2 处剩余的服务量来求取.

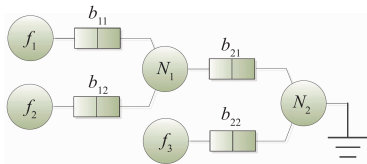


图4 多流-多节点模型

带宽分配分析法适用于时分复用的仲裁方式. 它仅考虑本路流的分配带宽,因此相对独立于其他路干扰流的注入情况,在带宽占用较大的情况下可以较准确的对延迟上界给出预期.

理论上,无论哪一种分析方法所得的延迟上界,都可证明为实际延迟的一个上界;另一方面,它们又都有可能引入悲观的延迟预测. 因此,对于一些给定仲裁方式的聚合模型,我们采用综合两种分析法的方式,得到一个更紧致、更准确的上界:

$$\bar{D} = \min\{h(\alpha_1, \hat{\beta}_{\text{sys_ip}}), h(\alpha_1, \hat{\beta}_{\text{sys_ip}})\}$$

4.3 多流-多节点模型

基于等价服务曲线分析方法,对多流-多节点的聚合模型求取目标流的端到端延迟上界时,具体流程如

下(假设 f_1 为目标流):

(1) 根据网络拓扑结构和流的注入情况,画出以 f_1 为主干的抽象干扰树图;

(2) 求取各节点对 f_1 的等价服务曲线: $\hat{\beta}_1^{N_1}, \hat{\beta}_1^{N_2}, \dots, \hat{\beta}_1^{N_n}$

(3) 求取整个系统对 f_1 的端到端等价服务曲线:

$$\hat{\beta}_{1_sys} = \hat{\beta}_1^{N_1} \otimes \hat{\beta}_1^{N_2} \otimes \dots \otimes \hat{\beta}_1^{N_n}$$

(4) 得出 f_1 的端到端延迟上界解析式:

$$\bar{D}_1 = h(\alpha_1, \hat{\beta}_{1_sys})$$

5 分析方法的实验评估

5.1 实验目的和设置

5.1.1 实验目的

本节中的实验主要目的有二. 一是对上文提到的几种理论分析方法所得的延迟上界的质量进行验证和评估. 为此,我们定义了紧致性 ξ 的概念:

$$\xi = \frac{D_{sim}}{D_{cal}} \times 100\%$$

其中 D_{sim} 表示在一次实际的仿真中观测到的最长延迟, D_{cal} 表示理论的延迟上界. 由于仿真实验中最长延迟能否达到延迟上界是一个概率事件,因此我们实验的另一个目的就是紧致性受到各个流、节点的哪些配置参数的影响进行讨论,这有助于我们理解在何时、何种配置情况下,网络会出现最坏情况的延迟和拥塞. 这为设计者考虑系统瓶颈提供了非常重要的参考.

5.1.2 仿真库与实验设置

我们在 SystemC 环境下建立一个时钟精确的仿真库来模拟片上的流和节点的行为. 仿真库中的重要基本元件包括源、阱、延迟-速率服务节点、传输延迟、FIFO 缓存等模块. 在我们的实验中,所有仲裁器均被抽象为持续工作的节点 $\beta_{1,0}$.

以图 3、4 中结构为实验对象,在全部实验中,到达曲线的参数范围为: $r \in (0, 1)$, $b \in [1, 128]$. 另外,考虑物理链路的带宽限制,TSPEC 模型中 $M = 1$, $p = 1$. 所有聚合流的实验均以加权轮询的仲裁为例进行分析,假设图 3 中 β_{σ} 节点的权重值为 $\phi_1 = \phi_2 = 1$,图 4 中 N_1 节点权重值 $\phi_1 = \phi_2 = 1$, N_2 节点的权重值 $\phi_{12} = 2$, $\phi_3 = 1$. 在图 7、8 的实验中,为了讨论阱的 R 和 T 对紧致性的影响,它们被作为实验的参数变量进行设置;而在其他实验中,阱均被抽象为一个服务能力为 $\beta_{1,0}$ 的节点. 每次仿真中每路流都会由源发出至少 5000 个数据包,经由服务节点最终到达阱.

5.2 目标流到达曲线对紧致性的影响

5.2.1 平均注入速率

实验中,仲裁器给目标流和干扰流各分配了 $R = 0.5$ 的最低服务保证. 目标流 $f_1 = r_1 t + 16$, $r_1 \in [0.05,$

0.5)], 干扰流 $f_2 = 0.5t + 32$. 本小节中专注讨论不同到达曲线模型的差别, 因此计算等价服务曲线时统一使用带宽分配法, 结果见图 5.

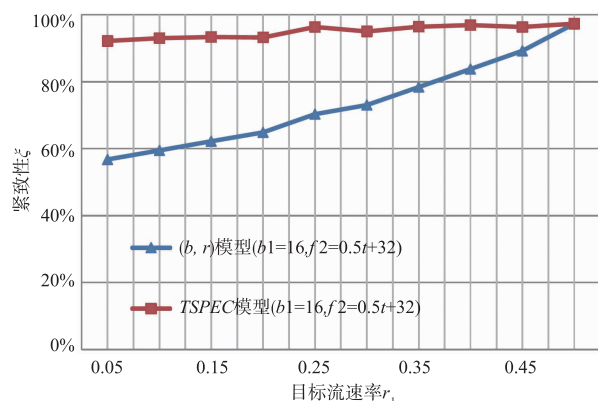


图5 (b, r) vs. TSPEC

(1) 从紧致性角度, TSPEC 模型的结果总是等于或优于 (b, r) 模型. 在相同参数配置下, r_1 和 R_{sys_ip} 的差值越大, TSPEC 模型的理论结果就越优于 (b, r) 模型.

(2) TSPEC 模型给出了非常稳定的紧致性, 趋近于 100%. ξ 曲线的走势说明 TSPEC 模型可很好的限制目标流平均速率 r_1 的改变对紧致性的影响; 而 (b, r) 模型的紧致性随 r_1 的增大而升高. 当 $r_1 = R_{sys_ip}$ 时, 两种模型的 ξ 曲线相交, 即随着 r_1 数值接近 R_{sys_ip} , (b, r) 模型的预测趋近于 TSPEC 模型的预测结果.

鉴于 TSPEC 模型总能给出更紧致结果, 在下文的单节点实验中, 到达曲线均采用 TSPEC 模型描述.

5.2.2 突发传输量

实验中目标流 $f_1 = r_1 t + b_1$, 其中 b_1 为 x 轴变量, 且 $b_1 \in [2^0, 2^7]$, 干扰流恒定为 $f_2 = 0.5t + 64$. 在三组实验中, 目标流 r_1 所占其配额带宽比例不同, 分别为 $r_1 = 0.2R_{sys_ip}/0.6R_{sys_ip}/R_{sys_ip}$. 若要理论延迟上界出现, 则一个数据包必须要在仲裁节点和阱节点都遇到最长的等候延迟, 并且遇到这个数据包的前一个包所经历的完整传输处理时间. 由图 6 实验结果可知:

(1) 随着目标流 b_1 增大, 结果的紧致性增高. 因为 b_1 数值很大时, 数据包在服务节点前的缓存中会产生更多的拥塞和等待, 这就给目标数据包等候前一个包的完整传输处理时间创造了条件.

(2) 总体上看, $r_1 = 0.5$ 的结果曲线优于 $r_1 = 0.3$ 的曲线, 而 $r_1 = 0.3$ 的曲线优于 $r_1 = 0.1$ 的曲线. 这进一步证明了当目标流占用其配额带宽值越多时, 目标流越容易遇到延迟上界.

5.3 服务曲线 $\beta_{R,T}$ 对紧致性的影响

当两路流在一个服务节点聚合时, 每路流所得的等价服务曲线 β_i 不仅取决于节点本身的服务曲线 $\beta_{R,T}$, 还取决于聚合种类和干扰流的注入情况. 为了评估服

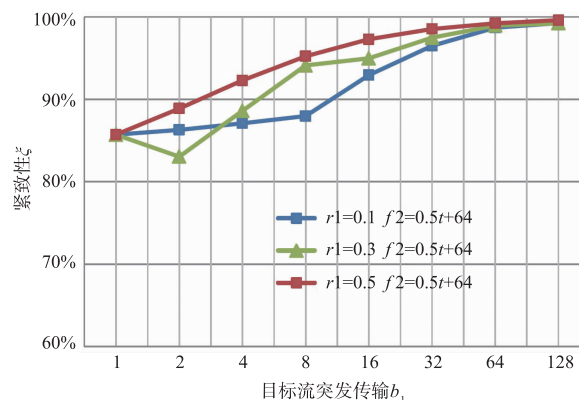


图6 目标流影响

务曲线对延迟上界紧致性的影响, 我们在本节中以单流-单节点实验为例进行讨论.

5.3.1 服务速率 R

实验中源发送的信息流 $f = 0.1t + b$, 节点服务曲线为 $\beta_{R,T}$, $R \in [0.3, 0.9]$. 当信息流的突发传输量 b 为 16/16/32 时, 延迟 T 分别等于 30/30/40 时钟周期. 如图 7, 服务速率 R 对延迟上界紧致性影响的趋势和流的平均速率 r 对紧致性的影响趋势相同, r 的数值越接近 R , 网络演算对延迟上界的预测越准确.

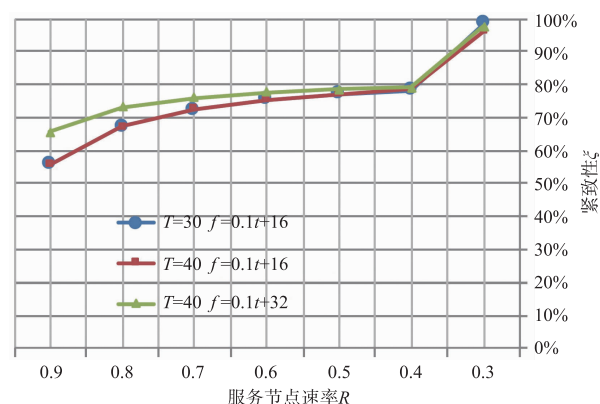


图7 服务曲线影响 (1)

5.3.2 服务延迟 T

实验结果如图 8. 信息流为 $f = 0.9t + b$, 四组实验中 b 的取值为 2/16/32/16, 服务速率 $R = 0.9$ 或 $R = 1$ 单位/周期, 服务延迟 T 取值以指数增长. 可观察到数值较大的 T 有助于最坏情况的出现, 但并非必要条件. 在一些实验中 (例如图中橙色的数据点) 即使在 $T = 0$ 的情况下, 依然可以得到 100% 准确的延迟上界预测. 另外, 最紧致的点也不一定出现在 T 取值最大处. 实际上, 对紧致性起到影响作用的并不是 T 本身的取值, 而是与信息流突发传输量的比值. 当时, 在数据流被服务之前它会在等候队列中缓存, 因而在得到服务时容易积累得到最大的突发传输量, 从而出现最长的等候时间.

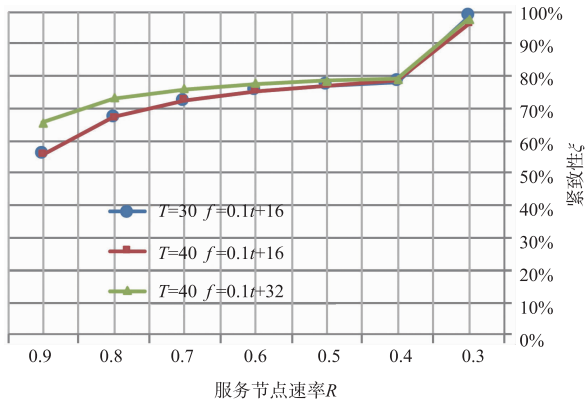


图8 服务曲线影响(2)

5.4 干扰流到达曲线对紧致性的影响

5.4.1 目标流负载较大时

在这两组实验中,目标流 $f_1 = 0.5t + 64$ 相较于于干扰流负载较大,干扰流设定为 $f_2 = r_2t + 8$,其中 $r_2 \in [0.1, 0.5]$,如图9中的紫色曲线(LP分析法)或绿色曲线(IP分析法);或对于 $f_2 = 0.1t + b_2$,其中 $b_2 \in [2^0, 2^7]$,如图10中的紫色曲线(LP分析法)或绿色曲线(IP分析法).当 r_2 或 b_2 增大时,IP分析法所得延迟上界的紧致性增高,而LP分析法的紧致性并无显著变化.当 r_2 或 b_2 取值较小时, ξ_{isp_lp} 远高于 ξ_{isp_ip} ,但随着 r_2 或 b_2 的增大,他们之间的差值逐渐缩小,直至 ξ_{isp_ip} 的值超过 ξ_{isp_lp} . IP分析法给出了目标流所能被保证的最小服务速率,它独立于干扰流的配置,行为模式而存在;而另一方面,LP分析法通过从总的服务能力中分离出干扰流所要求的服务,从而给出了目标流所能得到的剩余最大服务速率,这种方法考虑了当 f_2 没有使用完它的配额带宽,而 f_1 恰好占用了这一部分空闲带宽的情况.这解释了为何在 f_1 相对空闲时,LP分析法更准确,而当 f_2 的负载量与 f_1 相接近时,IP分析法所得延迟上界更贴近实际观测值.

5.4.2 目标流负载较小时

在这两组实验中,目标流 $f_1 = 0.1t + 64$ 相较于于干扰流负载较小,干扰流设定为 $f_2 = r_2t + 8$,其中 $r_2 \in [0.1, 0.9]$,如图9中的红色曲线(LP分析法)或蓝色曲线(IP分析法);或对于 $f_2 = 0.1t + b_2$,其中 $b_2 \in [2^0, 2^7]$,如图10中的红色曲线(LP分析法)或蓝色曲线(IP分析法). ξ_{isp_lp} 和 ξ_{isp_ip} 的变化趋势,以及它们之间的位置关系与前两组实验基本相同.值得关注的是,在某个特定的数据点右侧, ξ_{isp_ip} 接近于饱和值100%.图9中,这个点出现在 $b_2 = b_1 = 64$ 处,这是由于在此数据配置下,干扰流已经占满了所有它所得的配额带宽,目标流 f_1 并不能得到额外的空闲带宽.图10中,这个点出现在处.因为当 b_2 的取值增大到一定程度时,它会在短时间内对 f_1 产生强大的阻塞作用,使得

f_1 达到延迟上界.

5.4.3 多流-多节点的紧致性实验

当聚合节点的数量增加时,不仅计算端到端延迟上界的复杂度大大增高,实验中可配置的信息流、服务节点的参数也将级数增长,这极大的增加了搜索空间的范围和分析紧致性问题的难度.同时,干扰流的数量、聚合节点位置关系对紧致性的影响也是一个十分复杂的问题.因而在本文的实验中,对多流-多节点的聚合模型,我们仅初步讨论目标流的到达曲线模型对紧致性的影响.

在多流-单节点实验中,我们得到结论 TSPEC 模型总能比 (b, r) 模型得到更紧致的延迟上界.然而随着信息流数量的增多,对每路流都使用 TSPEC 模型,会导致 \bar{D} 的推算过程和最终得到的解析式十分复杂.因此,若我们能证明在多流-多节点实验中 (b, r) 模型同样能给出紧致的预测,那么将会大大降低分析模型的复杂程度.

图11描绘了一组3流-2节点的实验结果,目标流设定为 $f_1 = r_1t + 4$,而 N_1, N_2 、节点加入的干扰流分别设定为 $f_2 = r_2t + b_2, f_3 = r_3t + b_3$,其中, $r_2 = r_3 = 0.05/0.1, b_2 = b_3 = 4/16$. N_1, N_2 处加权轮询的权重配比分别为 $(1:1)/(2:1)$.延迟上界的分析使用了带宽分配(处理链路共享)和剩余服务量(处理缓存共享)综合的方法.不同颜色的曲线代表了干扰流恒定在不同配置上时目标流的 r_1 对紧致性的影响.可以看到即便我们使用了简化的到达曲线模型,依然可以给出相对紧致的延迟上界预测(均大于80%).在一些参数配置下,还给出了100%准确的延迟预测,其走势也和多流-单节点模型中对紧致性影响的趋势大体相同.因为在多流聚合的模型中,每路流所得的配额带宽相对较小,整个系统链路的空闲带宽也较小,即目标流很容易符合 r_1 接近 $R_{sys,1}$ 的条件.因此在分析多流-多节点聚合模型时,使用 (b, r) 到达曲线模型,也能给出紧致的延迟上界分析.

6 结语

在本文中,基于网络演算理论,我们讨论了两种到达曲线的模型和两种聚合节点的等价服务曲线分析方法;在实验中测评了不同的流、节点配置参数,是如何影响延迟上界的紧致性的;给出了评估不同分析法质量的手段,进而回答了何时、何种参数配置情况下,会出现预测的最坏延迟情况这个问题.得到如下结论:(1)网络演算分析的延迟上界是可到达的,在实验中,能出现100%准确预测的情况;(2)到达曲线的 TSPEC 模型在紧致性方面优于 (b, r) 模型,却有着更高的计算复杂度,而 (b, r) 模型在趋近于等价服务速率 R_{sys} 时,也能给出同样准确的延迟上界预测,它更适合多节点复杂网络

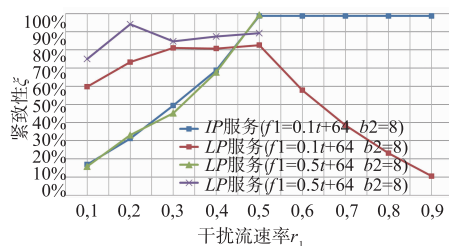


图9 干扰流影响 (1)

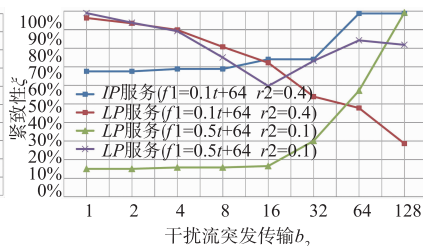


图10 干扰流影响 (2)

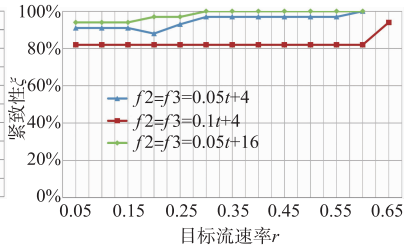


图11 多流-多节点实验

的分析建模; (3) 剩余服务量分析法适用于所有聚合模型, 而带宽分配法适用于时分多路复用的情况, 在片上常用的加权轮询仲裁方式中, 综合两种分析法可以得到更准确的延迟上界。

参考文献

- [1] 欧阳一鸣, 何鑫城, 等. 针对路径故障与局部拥塞的 NoC 容错路由算法[J]. 电子学报, 2016, 44(4): 920-925.
Ouyang Y, He X, et al. A fault-tolerant routing algorithm aiming at a path fault and local congestion in NoC[J]. Acta Electronica Sinica, 2016, 44(4): 920-925. (in Chinese)
- [2] 姚磊, 蔡觉平, 等. 基于内建自测技术的 Mesh 结构 NoC 无虚通道容错路由算法[J]. 电子学报, 2012, 40(5): 983-989.
Yao L, Cai J, et al. A fault-tolerant routing algorithm based on BIST for 2D-mesh Network-on-Chip without using virtual channels[J]. Acta Electronica Sinica, 2012, 40(5): 983-989. (in Chinese)
- [3] Xue S, et al. Parallel FFT implementation based on multi-core DSPs[A]. IEEE International Conference on Computational Problem-Solving[C]. IEEE, 2011. 426-430.
- [4] Cruz R L. A calculus for network delay, Part I: Network elements in isolation; Part II: Network analysis[J]. IEEE Transactions on Information Theory, 1991, 37(1): 114-141.
- [5] Agrawal R, Cruz R L, Okino C, et al. Performance bounds for flow control protocols[J]. IEEE/ACM Transactions on Networking, 1999, 7(3): 310-323.
- [6] LeBoudec J-Y, Thiran P. Network Calculus: A Theory of Deterministic Queuing Systems for the Internet: Vol 2050[M]. [S. l.]: Springer-Verlag, 2001.
- [7] Qian Y, Lu Z, Dou W. Analysis of worst-case delay bounds for best-effort communication in wormhole networks on chip[A]. 3rd ACM/IEEE International Symposium on Networks-on-Chip[C]. ACM/IEEE, 2009. 44-53.
- [8] Ogras U Y, Bogdan P, Marculescu R. An analytical approach for network-on-chip performance analysis[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2010, 29(12): 2001-2013.
- [9] Bogdan P, Marculescu R. Workload characterization and its impact on multicore platform design[A]. Proceedings of the Eighth IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis[C]. IEEE/ACM/IFIP, 2010. 231-240.
- [10] Wang J, Li Y, Peng Q. A novel analytical model for network-on-chip using semi-Markov process[J]. Advances in Electrical and Computer Engineering, 2011, 11(1): 111-118.
- [11] No. RFC 2210, The use of RSVP with IETF integrated services[S].

作者简介



龙彦辰 女, 1988 年生, 浙江大学电路与系统专业博士生. 研究方向为片上网络服务质量保证.

E-mail: yanchen@kth.se



沈海斌 男, 1967 年生, 浙江大学教授, 研究方向为面向智能安全的高性能计算及其芯片实现.

E-mail: shen_hb@zju.edu.cn



鲁中海 男, 1968 年生, 瑞典皇家工学院副教授, 研究方向为计算机体系结构.

E-mail: zhonghai@kth.se