

基于寄存器插入环的环形交换结构

Ring Fabric of Switch Based on Register Insert Ring

王一匡 周胜源 (桂林电子科技大学 信息与通信学院 广西 桂林 541004)

摘要: 弹性分组环(RPR)是基于分组业务优化的城域网络体系结构,其环形结构属于寄存器插入环。寄存器插入环同时具有时分复用和空分复用的特点,而且环形结构的调度简单,很适合作为交换结构使用。本文将寄存器插入环结构用于交换结构。以排队论为工具对环形交换结构进行建模,分析环形交换结构的性能。结果表明寄存器插入环结构是一种适合高速分组交换的结构。

关键词: 弹性分组环 环形交换结构 性能分析

1 引言

弹性分组环(RPR)是刚被 IEEE802.17 标准化的主要用于宽带 IP 光城域网的新型技术,也是业界第一次为基于分组业务而优化的城域网络体系结构,是城域网技术发展的重要方向。弹性分组环属于寄存器插入环,融合了共享媒体(时分交换)和空分交换的特性,很适合传输分组业务。由于环形结构本身的简单性,其调度问题很容易解决。因此我们尝试把弹性分组环的核心概念从物理结构中抽象出来尝试用于交换结构,提出一种新的基于寄存器出入环的环交换结构。

2 以RPR为代表的寄存器插入环

根据媒体访问的技术措施来划分,环网主要分为总线环(争用环)、令牌环(Token Ring)、时隙环(Slotted Ring)和寄存器插入环(Register/Buffer Insertion Ring)^[1]。前两种环形网络都属于共享媒体的时分结构,整个环的带宽作为一个整体,任意一个节点有业务进入环的时候,便将整个环锁定,其他端口的业务便不能进入,这样,整个环上所有节点的进入业务之和不会大于环的带宽,时隙环的媒体访问控制方式采用许多小的固定时隙实现,允许不止一个端口同时访问环,但业务在源节点剥离导致其空间重利用能力微弱,再加上应用较少,在这里不进行讨论。RPR 节点间采用点到点的逻辑连接方式连成环形,并

没有构成物理上的全连接,相邻节点间的环段带宽为所有端口所共享,可以实现带宽统计复用,这是时分(共享媒体)的特点。但是 RPR 共有 N (节点数)个环段总带宽共享,和传统的时分(共享媒体)相比具有很大的优势。同时由于分组只是顺着环在源和目的之间的端口上依次通过,在目的节点剥离,只使用源与目的节点间的线路,并不锁定整个环,其它节点间的通信可以同时进行,这样在空间上没有冲突的业务可以互不影响地利用各自线路的带宽,具有所谓的“空间重利用”能力,这也是空分交换的特点。通过“空间重利用”,比起前 2 种共享媒体环可以增加带宽(在极限情况下,最多可以到达 N 倍环段带宽),从而可以克服共享媒体(时分)交换的带宽限制缺陷。同时环形是一种简单的结构,每个节点都“知道”它们在一个环上,所以环上的分组不管从环上的那条路径传输都会最终到达它的目的地。由此,每个节点只需三种基本分组处理动作:输入本节点的插入分组、向下游节点转发途经的分组和“剥离”正确到达(目的地是节点)的分组,而不需要复杂调度甚至路由问题,从而可以使用更简单的协议,达到更高处理速度。综上所述,以 RPR 为代表的寄存器插入环一种是很适合高速分组传输的结构。我们以 RPR 为借鉴,将寄存器插入环结构用于交换结构,提出一种适合未来高速分组交换的环形交换结构。

① 基金项目:广西科学基金项目应用基础专项(桂科基 0832012);广西区研究生科创新计划(2008105950810M412)

收稿时间:2008-12-01

3 基于寄存器插入环的环形交换结构

端口采用两条传输方向相反的线路进行点对点的连接,从而形成与 RPR 相似的双环结构。在每一条环路方向上采用双转发缓存结构,一个缓存与用转发上游端口发送过来的分组(TB),另一缓存用于向环路发送本端口的分组(TX)。端口受到的分组先经过接受交换剥离本地接受的分组,需要转发的进入转发缓存(TB)。在 TB 和 TX 都有分组需要转发的情况下,优先发送 TB 中的分组(如图 1)。

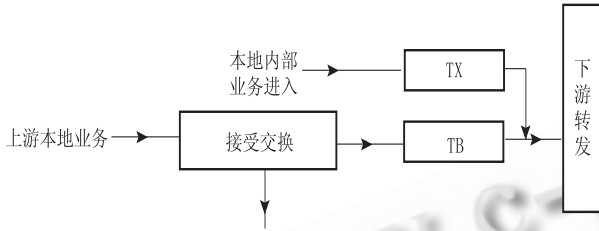


图 1 双缓存结构

由于环路的带宽为环段上所有的端口共享,为了满足每个端口的带宽需要就必须增加环段的带宽。对于交换结构来说通常端口带宽要求很高,对一条环路的带宽来说又不可能无限的宽,这就需要通过增加端口这间相连的环路数来解决。具体的办法我们会在下一部分详细的讲述,这里我们仍以双环为例。环形交换结构的工作方式和 RPR 相似,每个端口向上游端口发送分组在一条环路上进行向下游端口发送的分组在另一条传输方向相反的环路上进行。

4 环形结构的性能分析

基于寄存器插入环的交换结构是由若干个部端口组成的两个反向的传输环。分组由端口插入环路需要选择经过下游转发端口的个数较少的环进行传输在目的端口被剥离。如果在转发缓存(TB)和发送缓存(TX)同时有分组需要发送时优先发送转发缓存(TB)中的分组。转发缓存和发送缓存都按照先入先出(FCFS)的原则进行分组的排队发送。设端口数为 n ,对于这种属于寄存器插入环的环形型结构作以下假设:

- (1) 转发缓存和发送缓存的容量是无限大的。
- (2) 分组的服务时间和到达时间为独立的。
- (3) 各端口的时延为恒定值。
- (4) 端口发送和转发的分组的长度 L 具有相同的分布,且具有相同的一阶矩 \bar{L} 和二阶矩 \bar{L}^2 。设系统的

服务速度为 R ,系统的服务时间为 $B=L/R$ 且具有相同的一阶矩 \bar{B} 和二阶矩 \bar{B}^2 。

(5) 所有的端口具有相同的传输模式即端口 i 以概率 p_i 向端口 j 发送分组。

(6) 发送分组的到达的过程服从泊松分布,所有端口的分组平均到达速度相等都为 λ_{TX} 。对于转发缓存的分组输入过程也服从泊松分布。

这样环形结构性能分析模型与弹性分组环的分析模型相似可以看成是队头(HOL,Head Of Line)非强占的带优先级的 M/G/1 排队系统。[2,3]对于队头非强占的带优先级的 M/G/1 排队系统如果其优先级为 $k=1,2,\dots,K$ (优先级越高 k 值越小)那么优先级为 k 的业务等待时间为:

$$w_k = \frac{w_0}{(1 - \sum_{i=1}^{k-1} \rho_i)(1 - \sum_{i=1}^k \rho_i)} \quad (1)$$

其中, $w_0 = \frac{1}{2} \sum_{i=1}^K \lambda_i \bar{B}_i^2$ 即平均剩余服务时间(mean residual service time). ρ_i 为优先级为 i 的业务的吞吐量。则平均队长为

$$x_k = \lambda_k w_k + \rho_k \quad (2)$$

设 m 为分组讲过的转发缓存的平均数则

$$m = \sum_{j=2}^{N-1} (j-1) p_j \quad (3)$$

那么,转发缓存(TB)平均分组到达速度为

$$\lambda_{TB} = m \lambda_{TX} \quad (4)$$

在我们的内部端口组成的环形结构中有两种优先级($k=2$)的业务,一种为优先发送的转发缓存 TB 中的分组($k=1$)。一种为在转发缓存 TB 中没有分组的发送缓存 TX 的分组($k=2$)。对于每一个端口来说,业务的平均吞吐量为

$$\frac{\rho_i}{n} = \lambda_i \bar{B}_i \quad (5)$$

两种业务的吞吐量分别为

$$\rho_1 = \rho_{TB} = \lambda_{TX} m \bar{B} = m \frac{\rho_{TX}}{n} \quad (6)$$

$$\rho_2 = \frac{\rho_{TX}}{n} = \lambda_{TX} \bar{B} \quad (7)$$

于是可以得到转发缓存(TB)和发送缓存(TX)中的分组发送队列的平均队长。

$$x_1 = \lambda_{TX} m w_1 + \rho_1 = \lambda_{TX} m \frac{w_0}{(1 - \lambda_{TX} m B)} + \lambda_{TX} m \bar{B} \quad (8)$$

$$\begin{aligned} x_2 &= \lambda_{TX} w_2 + \rho_2 \\ &= \lambda_{TX} \frac{w_0}{(1 - \lambda_{TX} m B)(1 - \lambda_{TX} m B - \lambda_{TX} B)} + \lambda_{TX} \bar{B} \end{aligned} \quad (9)$$

转发缓存(TB)中分组的排队等待时间

$$w_{TB} = w_1 = \frac{w_0}{(1 - \lambda_{TX} m B)} \quad (10)$$

发送缓存(TX)中分组的排队等待时间为:

$$w_{TX} = w_2 + m w_{TB} \quad (11)$$

为了得到定性反映系统特性的近似结果,我们在相对简单的假设(均匀模式)下分析。假设端口的数目 n 为偶数(奇数的分析方法相同)且端口 i 发往端口 $j=i+N/2$ 的分组全部从考察环发送。端口 i 向端口 j 发送分组的概率为 $2/N$, 分组的服务时间为负指数分布并具有相同的一阶矩 $1/\mu$ 和二阶矩 $2/\mu^2$ 。那么分组需经过的转发缓存的平均个数为 $m=(n-2)/4$ 。环形结构的总的吞吐量最大值 $\rho_{MAX} = \frac{4N}{N+2}$ 在 $N \rightarrow \infty$ 时 $\rho_{MAX} = 4$ 。从上面对时延的推导可知 $1 - \lambda_{TX} (m+1) \bar{B} > 0$, 当 \bar{B} 为 1 时 $\lambda_{TX} < \frac{1}{m+1}$, 这就意味着环路的带宽要为端口带宽的 $m+1$ 倍。为了解决这一问题就要增加环路的数目。若有 n 个端口那么环路的数目增加为 $2 \times \lfloor \frac{n+2}{4} \rfloor$, 每一个传输方向的环路数为 $\lfloor \frac{n+2}{4} \rfloor$ 。端口向一个方向传输的分组均匀的在这 $\lfloor \frac{n+2}{4} \rfloor$ 条环路上发送, 当端口的分组到达速度为 1 时每一条环路上的端口分组到达速度为 $1/m+1$ 。这样可以保证每条环路上的分组达到速度不超过极限值。对于一个端口来说其吞吐量就为 1, 环形结构的总吞吐量就为 n 。

考察一个有三十个端口的环形结构, 每一个传输方向的环路数为 8, (共有 16 个环)(设 $1/\mu = 1$)其时延和队长性能如图 2(1)可以看出随着端口分组到达速度的增大, 分组的转发时延增加, 分组的发送时延增

加的更为剧烈。其队长也会激烈的增大。为了是改善环形结构的性能我们在每一个传输方向上增加两个环路(环路数为 10), 其时延性能的改善如图 2(c)(d)。从中我们可以看到其时延和队长性能的改善是十分明显的。环形结构的分组发送时延主要由分组平均经过的缓存数和分组的转发时延决定。增加环路数改善性能的实质是降低了分组的转发时延。还也以采用 RPR 结合 WDM 的弹性分组星环网^[4]扩展时使用的星环嵌套的方法减少分组经过的缓存数。如共有 N 个端口, 在进行嵌套时分为 K 组, 每组有 M 端口, 每一组的 M 个端口组成一个小环, 每个小环会有一个中心端口, 共有 K 个中心端口, 再在把这 K 个端口接成一个环结构, 小环上的端口需要与环外端口发送分组时由中心端口转发至目的端口所在的环, 这就所谓的星环嵌套。一个 64 端口的环状结构进行 8×8 嵌套后如图 3 原来发送的分组经过的转发缓存数最大为 31, 嵌套后发送的分组经过的转发缓存数最大 5。星环嵌套能明显的减少分组经过的转发缓存数, 大大降低降低分组的发送时延。其详细的性分析有待进一步研究。

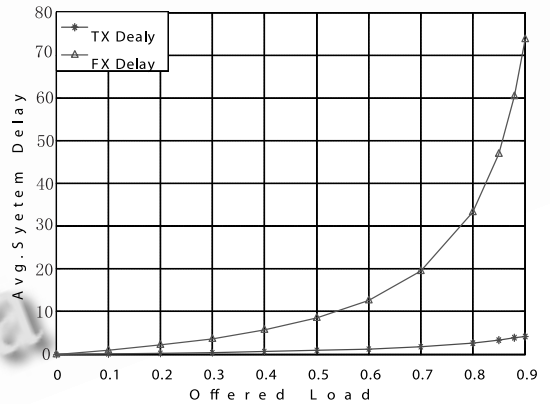


图 2 (a)端口时延性能

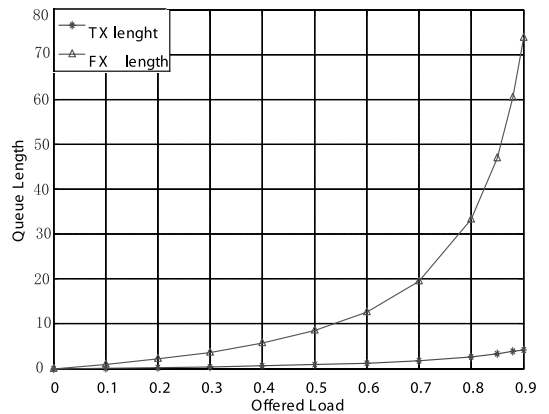


图 2 (b)端口队长性能

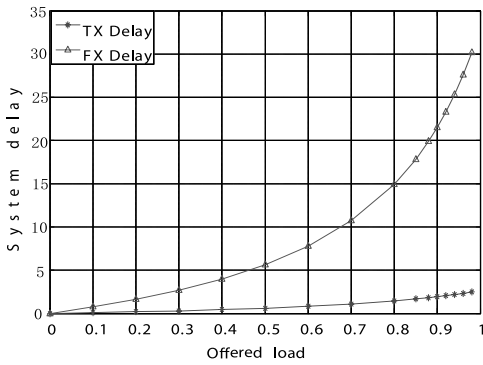


图 2 (c)端口时延性能

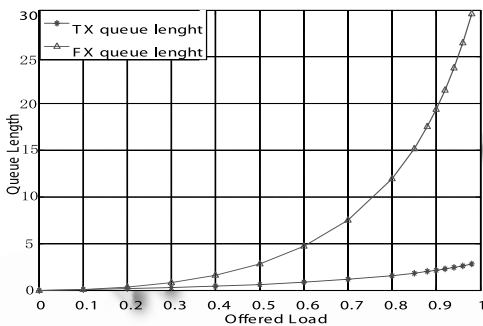


图 2 (d)端口队长性能

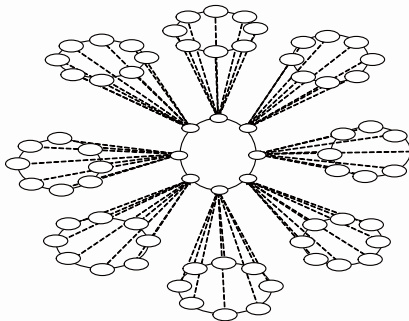


图 3 星环嵌套

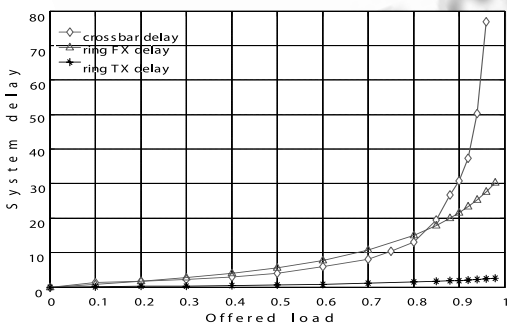


图 4 时延性能比较

5 与crossbar的性能比较

选取混合输入输出排队 CROSSBAR 结构做比较, Jeane S.-C. Chen, Thomas E. Stern 对混合输入输出

出排队结构做过建模分析。分析过程参见文献[5]。在均匀业务模式下比较两种结构的性能, 设每一个端口的分组等概的向其它端口发送分组。端口的分组到达速率 $\lambda \rightarrow 1$, crossbar 的加速度为 4 (crossbar 最大吞吐量可以达到 99%) 端口数为 30。在两种结构最大吞吐量都接近 100% 时, 其时延性能的对比如图 4 可以看出在分组到达速率小于 0.8 时两种结构的时延性能相差不多 crossbar 略好于环结构, 但在分组到达速率高于 0.8 时环形结构的时延性能要明显好于 crossbar 结构。所以可以得到这样的结论, 在均匀业务模式下寄存器插入环结构的排队时延性能要好于 crossbar 结构。

6 结论

本文在均匀业务模式下, 对寄存器插入环结构和 crossbar 结构的性能进行了分析比较, 根据前面的结果我们可以看出在均匀业务模式下环形结构的时延性能要好于 crossbar 结构。环形结构的复杂度要大大低于 crossbar 结构。可见寄存器插入环结构是一种适合高速分组交换的结构。

参考文献

- 1 Zhou SY, Wei MY, Cheng MS, Ao FL. An Observation on Resilient Packet StarRing. International Conference on Communications, Circuit and System (ICCCAS), Guilin, China, Jun 2006,3:1706 – 1707.
- 2 Zhou SY, Ao FL, Cheng MS, Delay and Throughput performance analysis of RPR’s three priorities traffic ICCAS, Chengdu, China, Jun2004,1:660 – 664.
- 3 Zhou SY, Ao FL, Cheng MS. RPR’s Triplex Transit Buffer Structure and Analysis of Its Delay Performance International Conference on Computer Communications (ICCC), Beijing, China, Sep 2004,1: 175 – 179.
- 4 姚洁,周胜源.弹性分组星-环的研究.光通信技术, 2007,8:26 – 28.
- 5 Chen JSC, Stern TE. Throughput Analysis, Optimal Buffer Allocation, and Traffic Imbalance Study of a Generic Nonblocking Packet Switch. IEEE Journal on Selected Aread in Communications, 1991,9(3):439 – 449.