

iSCSI 虚拟交换机包转发调度算法 FC-WFQ

韩 永,姚念民,蔡绍滨,陈金忠

(哈尔滨工程大学计算机科学与技术学院,黑龙江哈尔滨 150001)

摘 要: 与传统 IP 交换机不同,iSCSI 虚拟交换机不仅实现 IP 数据包的转发,还应完成 TCP、iSCSI 和 SCSI 的协议处理.短包优先调度算法可保证 iSCSI 控制及命令 PDU(Packet Data Unit)的优先传输,却未对该类数据包的转发带宽作出限制,该方法缺少针对 iSCSI 数据流特性的定量分析.通过排队理论建立 iSCSI 虚拟交换机转发 iSCSI PDU 过程的数学模型,然后提出一种 iSCSI 虚拟交换机的包调度算法 FC-WFQ(Flow Control-WFQ),最后应用网络仿真软件 ns-2 建立仿真场景对交换系统进行测试.随着命令到达率以及命令中读写比例的变化,FC-WFQ 对各数据流的转发带宽权重做相应的实时调整.实验结果显示,该调度算法可明显降低读写任务的平均响应时间,并显著提高 iSCSI 虚拟交换机的吞吐量.

关键词: SAN; SCSI; iSCSI 虚拟交换机; 调度算法; WFQ; 排队模型

中图分类号: TP393.05 **文献标识码:** A **文章编号:** 0372-2112(2013)03-0587-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2013.03.028

Packet Forwarding Scheduling Algorithm FC-WFQ for iSCSI Virtualization Switch

HAN Yong, YAO Nian-min, CAI Shao-bin, CHEN Jin-zhong

(College of Computer Science and Technology, Harbin Engineering University, Harbin, Heilongjiang 150001, China)

Abstract: Different from traditional IP switches, the iSCSI virtualization switch not only realizes the IP packets transmission, but also performs the processing of TCP, iSCSI and SCSI protocol. The short packet first scheduling algorithm insures the priority transmission of iSCSI command and status PDUs (Packet Data Unit), but does not impose restrictions on the forwarding bandwidth of such packets. Moreover, it is lack of quantitative analysis on iSCSI data flow characteristics for this method. The mathematical model of iSCSI PDUs forwarding process within the iSCSI virtualization switch was created by queuing theory. Then the packet scheduling algorithm called FC-WFQ (Flow Control-WFQ) was designed for the iSCSI virtualization switch used in the SANs. Finally the iSCSI virtualization switch with FC-WFQ was tested under certain network scenarios built by ns-2. In the experiments, while the command arrival rate and the ratio of read/write commands were changing, the FC-WFQ dynamically adjusted the forwarding bandwidth weight of each packet flow correspondingly. The results show that the scheduling algorithm can greatly reduce the mean response time of the read/write tasks, and significantly improve the throughput of the iSCSI virtualization switch.

Key words: SAN; SCSI; iSCSI virtualization switch; scheduling algorithm; WFQ; queuing model

1 引言

存储虚拟化技术屏蔽了不同存储设备的差异性并提供简单的访问方式,存在对称结构和非对称结构两种模式,这两种实现方式的根本区别在于数据流和控制流是否走相同的路径^[1].对称结构(带内方式)是指在储存设备和应用服务器的数据路径上实现存储设备的虚拟化,其数据和控制信息共用同一传输路径.

SAN^[2,3]将远程的存储设备和本地的用户连接起来,使用户感觉上是在接受本地存储设备的服务. iSC-

SI^[4]通过 IP 网络访问 SCSI 存储设备,因此可以用来构建廉价的存储区域网.为了设备的扩展性和虚拟化管理,在 iSCSI 存储区域网中建立虚拟交换机,管理分散的物理存储设备,以便为用户提供一个整体的逻辑存储空间^[5].

iSCSI 虚拟交换机是一种带内实现方式.由于控制流和数据流都需要通过交换机转发,其数据传输性能会受到影响.而且 iSCSI 虚拟交换机集成了交换和虚拟化功能,很容易成为系统的瓶颈.文献[6]通过缓存 iSCSI 会话 ID 结构的包转发方法来减少协议进程的开销,以

提高 iSCSI 虚拟交换系统的性能. 文献[7]在 iSCSI 虚拟交换机中增加中转同步调度模块, 从而加速读写任务的执行过程.

WFQ 的基础是 GPS 算法^[8]; 各种交换机调度算法在基于 iSCSI 数据流的性能评估中, WFQ 总带宽利用率最高, 且权重对于各个流的时延、时延抖动、速率、速率抖动的影响也得到了很好的体现^[9]. iSCSI 存储网络充斥着大量的 SCSI 协议的状态和控制命令的传输^[10], 由于 IP 网络的带宽限制和网络延迟, iSCSI initiator 和 iSCSI target 在确认对方发出的状态或控制短命令前将等待更多的时间. 文献[11]提出了一种抢占式短包优先调度算法; 文献[12]提出了 iSCSI 控制短命令在存储网络的交换机中优先传输的调度方法, 以改善任务平均响应时间和系统吞吐量, 其本质为具有两个优先级队列的 PQ 算法. 还有一些研究也是在传统 IP 网络交换机中进行的调度优化^[13,14].

然而, 与传统交换机不同, iSCSI 虚拟交换机不是简单地实现 IP 数据包的转发, 还应完成 TCP、iSCSI 和 SCSI 的协议处理, 以及逻辑地址和物理地址的相互转换. 因此, 在研究 iSCSI 虚拟交换机的包转发调度问题时, 还应考虑 iSCSI 的数据流特性以及交换系统的处理机能力.

本文使用排队论建立 iSCSI 虚拟交换机转发 iSCSI PDU 的数学模型, 并以此为基础设计一种 iSCSI 虚拟交换机的包调度算法 FC-WFQ. 该算法实时观测数据流特征, 为命令 PDU 和数据 PDU 动态分配转发带宽.

2 iSCSI 协议中的读写请求实现过程

下面结合文档 RFC3720^[15]分析 iSCSI 协议中 iSCSI 读写请求的具体实现. 如图 1 所示, iSCSI Read 实现过程比较简单. iSCSI initiator 发出 SCSI Read 请求以后, iSCSI target 接受请求并从存储设备中读出数据, 然后通过 Data-In 将数据返回给 iSCSI initiator. 如果数据量较大, 就要分成多个 Data-In 数据包进行传输, 并在最后一个数据包中带有结束标记. iSCSI initiator 请求的所有 Data-In 数据长度之和(即 iSCSI initiator 每次读请求数据的长度)不得超过设定的 MaxBurstLength 值.

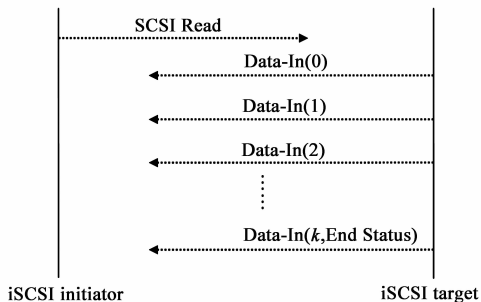


图1 iSCSI读请求过程的数据包交互流程

受几个协商关键字的影响, iSCSI Write 具有几种不同的实现方式. 以下介绍一种典型的 iSCSI Write 的实现方式(该方式允许传送非请求数据). 具体过程如图 2 所示: 首先 iSCSI initiator 发 SCSI 写命令, 并紧接着通过 Data-Out 的形式依次发送一部分非请求数据, iSCSI target 接收命令和数据. 如果仍有数据有待发送, iSCSI target 会发出一个 R2T 包, iSCSI initiator 根据 R2T 划定的范围继续发送数据. 重复上面的过程直到全部数据被 iSCSI target 接收. 最后 iSCSI target 返回 SCSI Response 告诉 iSCSI initiator 任务执行完成. R2T 的请求范围受 MaxBurstLength 值所限制.

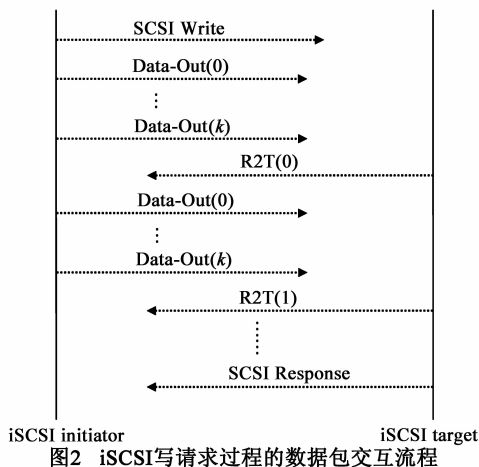


图2 iSCSI写请求过程的数据包交互流程

具有 iSCSI 虚拟交换机的存储系统如图 3 所示. 每当一个 iSCSI initiator 发送的 iSCSI 数据包到达虚拟交换系统, TCP、iSCSI 和 SCSI 几个协议对数据包进行处理以提取有效数据. 在将逻辑地址转化为物理地址后, 虚拟系统还要找到一条路径到达物理位置. 因此, 数据或读写请求的命令不得不转化为 SCSI command 有效载荷的形式, 数据将加载新的 iSCSI 协议头. 最终, 这个新产生的 iSCSI PDU 通过 IP 网络被传送到 iSCSI target. iSCSI target 回应的数据包将以同样的方式在虚拟交换系统中处理, 并最后返回到 iSCSI initiator.

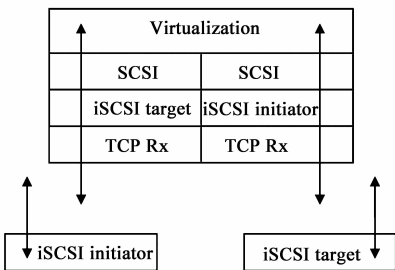


图3 iSCSI虚拟交换机内的包流动路径

文献[12]在交换机中优先转发 iSCSI 控制及命令 PDU, 实现了 iSCSI 任务的加速执行. 然而, 一个 iSCSI 会话中一般并发执行多个读写任务, 尤其当连接的 iSCSI

会话数量较多时, iSCSI 虚拟交换机要完成大量控制及命令 PDU 的协议处理和转发. 因此, 该方法虽然直观上保证了控制及命令 PDU 的转发高优先级, 却没有对其转发带宽进行节制, 从而抢占数据 PDU 的转发带宽, 则待处理的数据 PDU 由于长时间排队占据大量交换机缓存, 丢包率大幅提高; 而且随着数据 PDU 的转发速率的降低, 将引起恶性连锁反应.

如图 2 写请求的实现过程中, Write 命令包优先转发, 然而紧跟其后的 Data-Out 数据包的转发带宽无法获得保证, 则它们之后产生的 R2T(0) 控制包的到达延迟将增大, 由 R2T(0) 触发的 Data-Out 数据包的到达延迟也将增大且仍无法保证转发带宽, 如此反复恶化, 使得任务的响应时间增大以及整个交换系统的吞吐量降低. 因此, 系统应分别为控制命令 PDU 和数据 PDU 合理分配转发带宽.

3 iSCSI PDU 转发过程的排队模型

使用排队论对 iSCSI 虚拟交换机转发 iSCSI PDU 的过程进行分析, 图 4 表示这一过程的系统状态图.

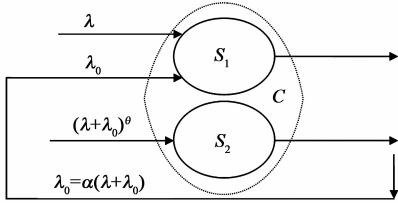


图4 iSCSI虚拟交换机转发iSCSI PDU的排队模型

控制及命令 PDU 的转发排队符合 [M/D/1] 模型^[16]. 读写命令包的到达率服从参数为 λ 的 Poisson 分布, R2T 控制包的到达率服从参数为 λ_0 的 Poisson 分布, S_1 为转发控制和命令 PDU 的服务台, 每个命令及控制 PDU 的转发服务时间为常数 $1/\mu_1$.

$\rho_1 = (\lambda + \lambda_0)/\mu_1$, 则控制和命令 PDU 的平均响应时间

$$W_1 = (2 - \rho_1)/2\mu_1(1 - \rho_1) \quad (1)$$

数据 PDU 的转发排队符合 [M^x/D/1] 模型^[16]. 在稳定状态下, S_1 转发控制和命令 PDU 的速率为 $(\lambda + \lambda_0)$. 若平均每个控制和命令 PDU 可触发 θ 个数据 PDU, 则数据包的成批到达率服从参数为 $(\lambda + \lambda_0)$ 的 Poisson 分布, 每批的数据包数量为常数 θ . S_2 为转发数据 PDU 的服务台, 每个数据包的转发服务时间为常数 $1/\mu_2$.

$\rho_2 = (\lambda + \lambda_0)\theta/\mu_2$, 则每批数据 PDU 的平均响应时间为数据 PDU 的平均等待时间 W_q 加上 θ 个数据 PDU 的服务时间 θ/μ_2 , 即

$$W_2 = W_q + \theta/\mu_2 = (1/\mu_2^2 + (\theta - 1)/\mu_2)/2(1 - \rho_2) + \theta/\mu_2 \quad (2)$$

经 S_2 转发的数据 PDU 中, 属于未完成的任务的数据包的比例为 α , 则触发的 R2T 控制包的到达率为 $\alpha(\lambda + \lambda_0)$, 即

$$\lambda_0 = \alpha(\lambda + \lambda_0) \quad (3)$$

经变形, 可得

$$\lambda_0 = \alpha\lambda/(1 - \alpha) \quad (4)$$

$$\lambda + \lambda_0 = \lambda/(1 - \alpha) \quad (5)$$

由式(4)可以看出, 初始读写命令中, 大数据写命令的比例越大, 则 α 越大; 若初始命令全部为读命令, 则 $\alpha = 0$.

控制和命令 PDU 在虚拟交换机内进行数据校验和内存拷贝的代价较低^[17]; 相比之下, 转发一个数据 PDU 要耗费的处理机资源为前者的 δ 倍. 因此, 若转发一个数据 PDU 需耗费的处理机资源为 ϵ , 则转发一个控制和命令 PDU 需耗费的处理机资源为 ϵ/δ .

C 表示 iSCSI 虚拟交换机的处理机, S_1 和 S_2 的转发能力都由 C 提供. 因此, C 的处理机能力为 S_1 和 S_2 的处理机能力之和. 若所有的处理机资源只用来转发数据 PDU, 且在单位时间内转发的数据 PDU 的数量为 P , 则单位时间内 C 的处理机能力可表示为 $P\epsilon$; S_1 和 S_2 在单位时间内可以转发的控制命令 PDU 和数据 PDU 的数量分别为 μ_1 和 μ_2 , 则单位时间内 S_1 和 S_2 的处理机能力分别表示为 $\mu_1\epsilon/\delta$ 和 $\mu_2\epsilon$, 则 $P\epsilon = \mu_1\epsilon/\delta + \mu_2\epsilon$, 即

$$P = \mu_1/\delta + \mu_2 \quad (6)$$

由转发模型可见, 每完成一个读写任务, 应平均转发 $(\lambda + \lambda_0)/\lambda$ 个控制及命令 PDU, 以及平均转发 $(\lambda + \lambda_0)\theta/\lambda$ 个数据 PDU. 则一个读写任务的平均响应时间

$$W = (W_1 + W_2)(\lambda + \lambda_0)/\lambda = (W_1 + W_2)/(1 - \alpha) \quad (7)$$

将式(6)代入式(7)可解得: 当

$$\mu_2 = (2P\delta\theta^2 - 2\alpha P\delta\theta^2 + \lambda\theta - \lambda\theta^2)/2(2 + \delta\theta^2 - \alpha\delta\theta^2 - \alpha) \quad (8)$$

时, W 可获得最小值.

4 FC-WFQ 算法

WFQ^[18] 是一种基于流的拥塞管理算法, 表示加权公平排队 (weighted fair queuing). 到达的数据包被分成多个流, 每个流都被分配给一个 FIFO 队列. WFQ 使用虚时间的概念. 调度器根据数据包的离开虚时间排序, 对数据包提供服务. 假设 N 是数据流队列的个数, 则 WFQ 的时间复杂度为 $O(N)$. 下面提出 iSCSI 虚拟交换机包转发调度算法 FC-WFQ (Flow Control-WFQ).

传统 IP 交换机仅完成数据转发功能, 因而转发数据包所需代价与该包长度成正比; 而 iSCSI 虚拟交换机还应完成 TCP、iSCSI 和 SCSI 的协议处理, 以及逻辑地址

和物理地址的相互转换,所以转发数据包所需代价不是简单地和数据包长度成正比,而是与各类协议处理和地址映射的操作有关.因此不再标记数据包的物理转发长度,而是以转发该包应耗费的处理机资源来标记其逻辑转发长度.

FC-WFQ 以 iSCSI PDU 为基本的调度转发单元.虚拟交换机转发一个数据 PDU 要耗费的处理机资源为转发一个控制和命令 PDU 的 δ 倍.设定数据 PDU 的逻辑转发长度为 l ,则控制和命令 PDU 的逻辑转发长度应为 l/δ .

式(8)中,记可使 W 获得最小值的 μ_2^* 为 μ_2^* ,则为两类 PDU 分配的最佳转发带宽之比 $B_1/B_2 = (P - \mu_2^*)/\mu_2^*$; 既在 FC-WFQ 算法中,两类 PDU 流的转发权重之比 $r_1/r_2 = B_1/B_2$.

每 Δt 时间段内统计初始读写命令包到达率 λ 以及 R2T 包到达率 λ_0 ,从而实时更新参数 α ;同时,每 Δt 时间段内统计数据包到达率,从而实时更新参数 θ .

综上,FC-WFQ 算法遵循基本的 WFQ 算法,可用如下伪代码说明:

```

1. /* Procedure of classifying PDU b to FIFO queue */
2.   If b is a Command&Control-PDU {
3.     Set FIFO_no = 0;
4.     Count_Comm&Contr-PDU ++;
5.   } Else {
6.     Set FIFO_no = 1;
7.     Count_Data-PDU ++;
8.   }
9.   t ++;
10.  PushFIFO ( b, FIFO_no );
11.  If t =  $\Delta t$ 
12.    AdjustWeight ( );
13.
14.  PushFIFO ( b, FIFO_no ) {
15.    If FIFO_no is 0 {
16.      Push b to FIFO-0;
17.      Set b_Length =  $l/\delta$ ;
18.      Set  $S_{0,b} = \max \{ F_{0,previous}, V(a_{0,b}) \}$ ;
19.      Set  $F_{0,b} = S_{0,b} + l/\delta r_0$ ;
20.    } Else {
21.      Push b to FIFO-1;
22.      Set b_Length = l;
23.      Set  $S_{1,b} = \max \{ F_{1,previous}, V(a_{1,b}) \}$ ;
24.      Set  $F_{1,b} = S_{1,b} + l/r_1$ ;
25.    }
26.  }
27.
28.  AdjustWeight ( ) {

```

```

29.    Set  $\lambda = \text{Command\&Control-PDU} + \text{Count\_Data-PDU}$ ;
30.    Set  $\theta = \text{Count\_Data-PDU}/\text{Command\&Control-PDU}$ ;
31.    Set  $\mu_2 = (2P\delta\theta^2 - 2aP\delta\theta^2 + \lambda\theta - \lambda\theta^2)/2(2 + \delta\theta^2 - \alpha\delta\theta^2 - \alpha)$ ;
32.    Set  $r_0/r_1 = (P - \mu_2)/\mu_2$ ;
33.    Set t = 0;
34.    Set Count_Comm&Contr-PDU = 0;
35.    Set Count_Data-PDU = 0;
36.  }

```

5 测试以及性能评估

应用网络仿真软件 ns-2 构造存储区域网络, iSCSI initiator 与 iSCSI 虚拟交换机之间的链路带宽和延迟为“1Gbps/100 μ s”; iSCSI 虚拟交换机与 iSCSI target 之间的链路参数与前者相同. iSCSI 虚拟交换机连接 3 个 initiator 和 3 个 target. 设定 iSCSI 虚拟交换机的转发带宽为 80MB/s, 在 ns-2 仿真中并不体现 FC-WFQ 算法复杂度造成的时延, 但在虚拟交换机仿真中会体现各类协议处理及地址转换造成的 iSCSI PDU 转发时延. 设定 iSCSI PDU 的最大数据长度为 8KB, MaxBurstLength 的值为 128KB.

本实验在以上构建的网络环境中完成 iSCSI 读写任务仿真过程. 设定 iSCSI 会话中的 TCP 连结数量为 5, 既执行命令的并发程度为 5. 每次读命令的请求数据大小为 128KB, 写命令的请求数据大小为 512KB, iSCSI 虚拟交换机收到来自 iSCSI initiator 的读写命令总到达率符合参数为 λ 的 Poisson 分布. 将文献[12]中的算法记为 TQ-PQ (Two Queue-PQ) 算法. 实验对 FCFS、WFQ^[18]、WRR^[19]、TQ-PQ^[12]、FC-WFQ 算法进行性能对比. WRR 和 WFQ 算法取不同权重测试结果中的最佳值.

图 5 和图 6 显示了当读写命令各占 50%, 随着初始读写命令到达率 λ 的增大, 读写任务平均响应时间和交换机吞吐量的变化. 随着初始命令到达率 λ 的提高, 读写任务平均响应时间和交换机吞吐量也都随之增大.

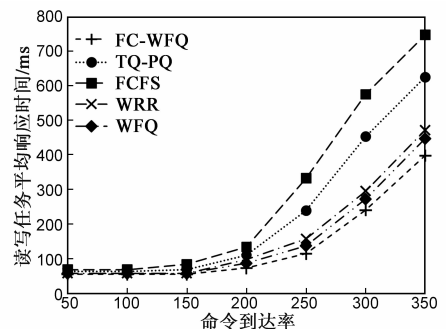


图5 读写命令各占50%时读写任务平均响应时间和命令到达率之间的关系

当 λ 较小时, 交换系统工作在轻负荷状态, 五种算

法下的读写任务平均响应时间和交换机吞吐量均没有明显差别;随着 λ 增大,交换系统负荷也逐渐增大,无法为两类 iSCSI PDU 流都提供绝对充足的转发带宽,此时对处理机资源的合理应用显得尤为重要,FC-WFQ 算法对转发带宽的合理分配优势得以体现。

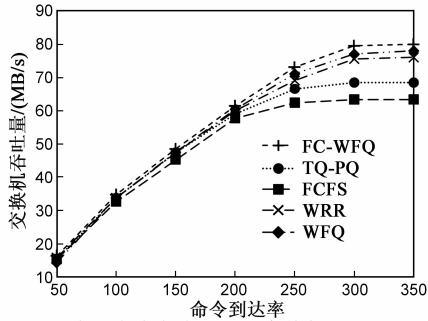


图6 读写命令各占50%时交换机吞吐量和命令到达率之间的关系

图 7 和图 8 显示了当命令到达率 λ 为 250,随着写命令比例的增大,五种算法下的读写任务平均响应时间和交换机吞吐量.由于 λ 较大,随着写比例的增大,交换系统负荷逐渐增大,直至工作在满负荷状态.读写任务平均响应时间和交换机吞吐量也都随之增大。

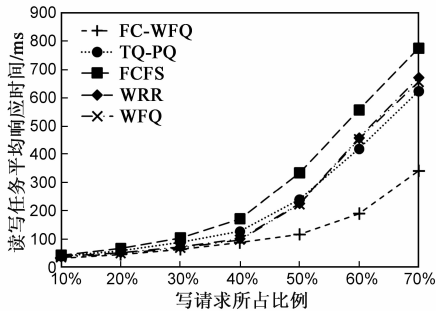


图7 命令到达率为250时读写任务平均响应时间和写请求所占比例之间的关系

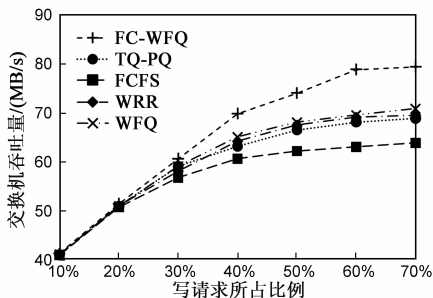


图8 命令到达率为250时交换机吞吐量和写请求所占比例之间的关系

初始读写命令到达率 λ 为定值,然而随着写命令比例的增大,R2T控制包的到达率 λ_0 也逐渐增大,则排队模型中的参数 α 和 θ 也随之变化.WRR 和 WFQ 算法中的转发带宽权重分配是固定的,而 FC-WFQ 算法根据这些 iSCSI 数据流的特征参数实时计算 μ_2^* ,动态地分

配转发带宽,从而实现最佳的交换系统性能.写命令比例较大时,FC-WFQ 算法下的读写任务平均响应时间和交换机吞吐量性能都得到了很大的改善。

6 总结

本文首先对 iSCSI 协议的通信特点进行分析,然后利用排队理论建立 iSCSI 虚拟交换机转发 iSCSI PDU 过程的数学模型,以此作为包转发调度算法 FC-WFQ 的理论基础.最后,应用网络仿真软件 ns-2 建立仿真场景对交换系统进行测试.实验结果显示,当交换系统工作在满负荷状态下,FC-WFQ 调度算法可明显降低读写任务的平均响应时间,并显著提高 iSCSI 虚拟交换机的吞吐量。

参考文献

- [1] 白勇,朱清新.基于 iSCSI 的 IP SAN 网络存储技术及其安全性研究[J].计算机研究与发展,2009,46(z1):12-17.
BAI Yong, Zhu Qing-xin. The iSCSI-based IP SAN network storage technology and its security research [J]. Journal of Computer Research and Development, 2009, 46(z1): 12-17. (in Chinese)
- [2] Barry Phillips. Have storage area networks come of age? [J]. Journal of IEEE Computer, 1998, 31(7): 10-12.
- [3] 谢长生,傅湘林,等.一种基于 iSCSI 的 SAN 的研究与实现[J].计算机研究与发展,2003,40(5):246-251.
XIE Chang-sheng, Fu Xiang-lin, et al. The study and implementation of a new iSCSI-based SAN [J]. Journal of Computer Research and Development, 2003, 40(5): 246-251. (in Chinese)
- [4] Meth K Z. Design of the iSCSI protocol [A]. Ben Kobler. Global Access to Distributed Storage [C]. DC, USA: IEEE Computer Society, 2003. 116-122.
- [5] David T. Volume managers in Linux [EB/OL]. http://www.sistina.com, 2001-04-20/2011-12-10.
- [6] Chung Yi-cheng. A packet forwarding method for the iSCSI virtualization switch [A]. Zhu Yi-feng. The 4th International Workshop on Storage Network Architecture and Parallel I/Os' 07 [C]. DC, USA: IEEE Computer Society, 2007. 35-40.
- [7] HAN Yong. Transfer-sycro-control mechanism of iSCSI virtualization switch to speed up its read/write requests [A]. Ni Jun. The 5th International Conference on Internet Computing for Science and Engineering' 10 [C]. DC, USA: IEEE Computer Society, 2010. 139-144.
- [8] Abhay K P, Robert G G. A generalized processor sharing approach to flow control in integrated services networks: the single-node case [J]. Journal of IEEE/ACM Trans on Networking, 1993, 1(3): 344-357.
- [9] 牛智,杨建军.基于 iSCSI 协议的调度算法性能比较[J].

- 江南大学学报(自然科学版), 2008, 7(5): 521 – 525. (in Chinese)
- NIU Zhi, YANG Jian-jun. Performance evaluation of different scheduling algorithms with iSCSI traffic[J]. Journal of Jiangnan University (Natural Science Edition), 2008, 7(5): 521 – 525. (in Chinese)
- [10] Intel Corp. Intel iSCSI Project[EB/OL]. <http://sourceforge.net/projects/intel-iscsi>, 2001-03-07/2011-12-30.
- [11] 李文杰, 刘斌. 输入排队中抢占式的短包优先调度算法[J]. 电子学报, 2005, 33(4): 577 – 583.
LI Wen-jie, LIU Bin. Preemptive short-packet-first scheduling in input queuing switches[J]. Acta Electronica Sinica, 2005, 33(4): 577 – 583. (in Chinese)
- [12] YAO Nian-min. Optimization of the switches in storage networks[A]. Shi Yong. The 7th International Conference on Computational Science '07[C]. Berlin, Heidelberg: Springer-Verlag, 2007. 66 – 73.
- [13] 马祥杰, 兰巨龙, 等. 输入排队 Crossbar 架构下的流量模型[J]. 电子学报, 2009, 37(1): 170 – 174.
MA Xiang-jie, LAN Ju-long, et al. Traffic model for input-queued crossbar fabric[J]. Acta Electronica Sinica, 2009, 37(1): 170 – 174. (in Chinese)
- [14] 涂晓东, 李乐民. 一类基于调度表的公平轮循调度算法[J]. 电子学报, 2001, 29(9): 1290 – 1293.
TU Xiao-dong, LI Le-min. A kind of schedule table based fair round robin scheduling algorithms[J]. Acta Electronica Sinica, 2001, 29(9): 1290 – 1293. (in Chinese)
- [15] Satran J. Internet Small Computer Systems Interface (iSCSI)[EB/OL]. <http://www.ietf.org/rfc/rfc3720.txt>, 2005-09-20/2011-02-30.
- [16] 盛友招. 排队论及其在计算机通信中的应用[M]. 北京: 北京邮电大学出版社, 1998.
- [17] Abhijeet J. A scalable and high performance software iSCSI implementation[A]. Jim Gray. The 4th USENIX Conference on File and Storage Technologies[C]. CA, USA: USENIX Association, 2005. 20 – 23.
- [18] Stiliadis D, Varma A. Latency-rate servers: A general model for analysis of traffic scheduling algorithms[J]. IEEE/ACM Transactions on Networking, 1998, 6(5): 611 – 624.
- [19] Shimonishi H, Yoshida M. An improvement of weighted round robin cell scheduling in ATM networks[J]. IEEE GLOBE-COM, 1997, 23(2): 1119 – 1123.

作者简介



韩永男, 1982年4月生于山东省枣庄市. 博士研究生, 主要研究方向为网络存储的性能与安全.

E-mail: hanyong@hrbeu.edu.cn



姚念民男, 1974年9月生于黑龙江省大庆市. 教授, 博士生导师. 主要研究方向为网络存储, 无线传感器网络等.

E-mail: yaonianmin@hrbeu.edu.cn