**调整Linux I/O调度器优化系统性能**

2017.09.02 01:56:32 来源:[51cto](http://www.techweb.com.cn/network/system/2017-09-02/2581460.shtml) 作者:51cto ( [条评论](http://www.techweb.com.cn/network/system/2017-09-02/2581460.shtml#wypl) )

前言

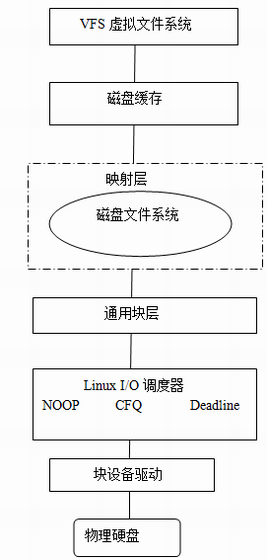
Linux I/O 调度器是Linux内核中的一个组成部分，用户可以通过调整这个调度器来优化系统性能。本文首先介绍Linux I/O 调度器的结构，然后介绍如何根据不同的存储器来设置Linux I/O 调度器从而达到优化系统性能。



Linux I/O 系统简介

Linux I/O调度器(Linux I/O Scheduler)是LinuxI/O体系的一个组件，它介于通用块层和块设备驱动程序之间。如图 1 所示。

图1 Linux I/O调度器介于通用块层和块设备驱动程序之间



当Linux内核组件要读写一些数据时，并不是请求一发出，内核便立即执行该请求，而是将其推迟执行。当传输一个新数据块时，内核需要检查它能否通过。Linux IO调度程序是介于通用块层和块设备驱动程序之间，所以它接收来自通用块层的请求，试图合并请求，并找到最合适的请求下发到块设备驱动程序中。之后块设备驱动程序会调用一个函数来响应这个请求。

Linux整体I/O体系可以分为七层，它们分别是：

VFS虚拟文件系统：内核要跟多种文件系统打交道，内核抽象了这VFS，专门用来适配各种文件系统，并对外提供统一操作接口。 磁盘缓存：磁盘缓存是一种将磁盘上的一些数据保留着RAM中的软件机制，这使得对这部分数据的访问可以得到更快的响应。磁盘缓存在Linux中有三种类型：Dentry cache ，Page cache ， Buffer cache。 映射层：内核从块设备上读取数据，这样内核就必须确定数据在物理设备上的位置，这由映射层(Mapping Layer)来完成。 通用块层：由于绝大多数情况的I/O操作是跟块设备打交道，所以Linux在此提供了一个类似vfs层的块设备操作抽象层。下层对接各种不同属性的块设备，对上提供统一的Block IO请求标准。 I/O调度层：大多数的块设备都是磁盘设备，所以有必要根据这类设备的特点以及应用特点来设置一些不同的调度器。 块设备驱动：块设备驱动对外提供高级的设备操作接口。 物理硬盘：这层就是具体的物理设备。

5种类型的Linux I/O调度器

Linux 从2.4内核开始支持I/O调度器,到目前为止有5种类型：Linux 2.4内核的 Linus Elevator、Linux 2.6内核的 Deadline、 Anticipatory、 CFQ、 Noop，其中Anticipatory从Linux 2.6.33版本后被删除了。目前主流的Linux发行版本使用Deadline、 CFQ、 Noop三种I/O调度器。下面依次简单介绍：

1 Linus Elevator

在2.4 内核中它是第一种I/O调度器。它的主要作用是为每个设备维护一个查询请求，当内核收到一个新请求时，如果能合并就合并。如果不能合并，就会尝试排序。如果既不能合并，也没有合适的位置插入，就放到请求队列的最后。

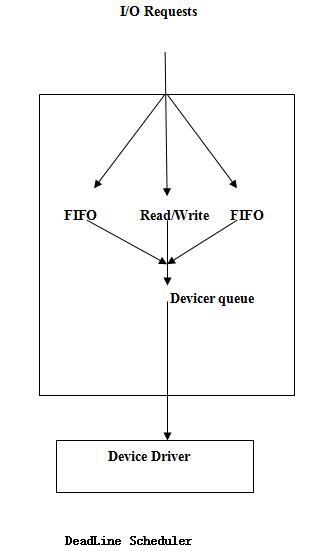
2 Anticipatory

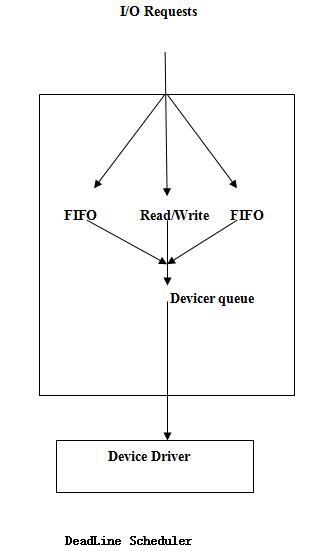
Anticipatory的中文含义是"预料的，预想的"，顾名思义有个I/O发生的时候，如果又有进程请求I/O操作，则将产生一个默认的6毫秒猜测时间，猜测下一个进程请求I/O是要干什么的。这个I/O调度器对读操作优化服务时间，在提供一个I/O的时候进行短时间等待，使进程能够提交到另外的I/O。Anticipatory算法从Linux 2.6.33版本后被删除了，因为使用CFQ通过配置也能达到Anticipatory的效果。

3 DeadLine

Deadline翻译成中文是截止时间调度器，是对Linus Elevator的一种改进，它避免有些请求太长时间不能被处理。另外可以区分对待读操作和写操作。DEADLINE额外分别为读I/O和写I/O提供了FIFO队列。Deadline的工作流程如图 2 所示。

图2 Deadline的工作流程

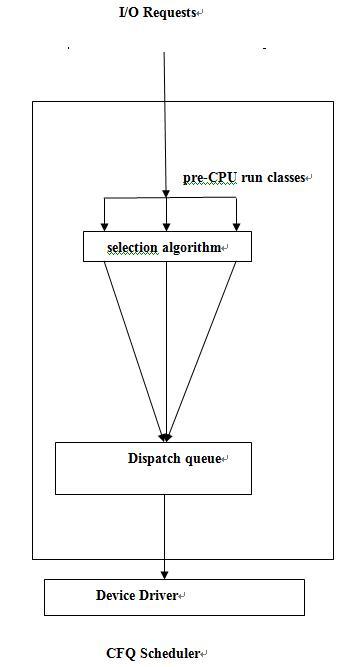


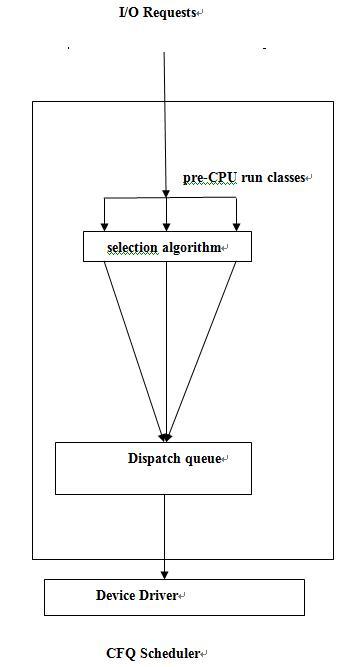


4 CFQ

CFQ全称Completely Fair Scheduler ，中文名称完全公平调度器，它是现在许多 Linux 发行版的默认调度器，CFQ是内核默认选择的I/O调度器。它将由进程提交的同步请求放到多个进程队列中，然后为每个队列分配时间片以访问磁盘。对于通用的服务器是最好的选择,CFQ均匀地分布对I/O带宽的访问。CFQ为每个进程和线程,单独创建一个队列来管理该进程所产生的请求,以此来保证每个进程都能被很好的分配到I/O带宽，I/O调度器每次执行一个进程的4次请求。该算法的特点是按照I/O请求的地址进行排序，而不是按照先来后到的顺序来进行响应。简单来说就是给所有同步进程分配时间片，然后才排队访问磁盘，CFQ的工作流程如图 3 所示 。

图3 CFQ的工作流程





5 NOOP

NOOP全称No Operation,中文名称电梯式调度器，该算法实现了最简单的FIFO队列，所有I/O请求大致按照先来后到的顺序进行操作。NOOP实现了一个简单的FIFO队列,它像电梯的工作主法一样对I/O请求进行组织。它是基于先入先出(FIFO)队列概念的 Linux 内核里最简单的I/O 调度器。此调度程序最适合于固态硬盘。NOOP的工作流程如图4 所示。

图4 NOOP的工作流程

I/O调度器的选择

目前主流Linux发行版本使用三种I/O调度器：DeadLine、CFQ、NOOP，通常来说Deadline适用于大多数环境,特别是写入较多的文件服务器，从原理上看，DeadLine是一种以提高机械硬盘吞吐量为思考出发点的调度算法，尽量保证在有I/O请求达到最终期限的时候进行调度，非常适合业务比较单一并且I/O压力比较重的业务，比如Web服务器，数据库应用等。CFQ 为所有进程分配等量的带宽,适用于有大量进程的多用户系统，CFQ是一种比较通用的调度算法，它是一种以进程为出发点考虑的调度算法，保证大家尽量公平,为所有进程分配等量的带宽,适合于桌面多任务及多媒体应用。NOOP 对于闪存设备和嵌入式系统是最好的选择。对于固态硬盘来说使用NOOP是最好的，DeadLine次之，而CFQ效率最低。

查看Linux系统的 I/O调度器

查看Linux系统的I/O调度器一般分成两个部分，一个是查看Linux系统整体使用的I/O调度器，另一个是查看某磁盘使用的I/O调度器。

查看当前系统支持的I/O调度器，使用如下命令：

清单 1. 查看当前系统支持的I/O调度器

# dmesg | grep -i scheduler [    1.508820] io scheduler noop registered [    1.508827] io scheduler deadline registered [    1.508850] io scheduler cfq registered (default)

清单1的代码显示cfq是目前的I/O调度器。

查看某块硬盘的IO调度算法I/O调度器，使用如下命令：

清单2. 查看一个硬盘使用的I/O调度器

# cat /sys/block/sda/queue/scheduler noop deadline [cfq]

清单2显示当前使用的调度器是cfq，就是括号括起来的那一个。

修改Linux系统的 I/O调度器

修改Linux系统的 I/O调度器有三种方法，分别是使用shell命令、使用grubby命令或者修改grub配置文件。下面依次介绍：

使用shell命令

Linux下更改的I/O调度器很简单。不需要更新内核，可以使用shell命令修改：

清单3. 查使用shell命令

#echo noop > /sys/block/sdb/queue/scheduler

清单3的命令把noop设置为一个磁盘的I/O调度器，你可以随时更改而无需重启计算机。

永久修改默认的I/O调度器

使用shell命令修改I/O调度器，只是临时修改，系统重启后，修改的调度器就会失效，要想修改默认的调度器，有两种方法使用grubby命令或者直接编辑grub配置文件。

使用grubby命令

例如需要把I/O调度器从cfq调整成 DeadLine ，命令如下：

清单4.使用grubby命令

# grubby --grub --update-kernel=ALL --args="elevator=deadline"

清单4的命令，通过设置内核加载参数, 这样当机器重启的时候，系统自动把所有设备的 I/O调度器变成 DeadLine 。

使用编辑器修改配置文件

也可以直接编辑grub的配置文件 ，通过修改grub配置文件，系统自动把所有设备的 I/O调度器变成cfq。操作过程如下：

清单5 使用vi编辑器修改grub配置文件

#vi cat /etc/default/grub #修改第五行，在行尾添加# elevator= cfq  然后保存文件，重新编译配置文件， #grub2-mkconfig -o /boot/grub2/grub.cfg

重新启动计算机系统即可。

总结

Linux I/O调度器是 Linux 内核中的一个组成部分，用户可以通过根据不同的存储器来设置 Linux I/O 调度器从而达到优化系统性能。 一般来说 NOOP 调度器最适合于固态硬盘，DeadLine 调度器适用于写入较多的文件服务器，比如Web服务器，数据库应用等，而CFQ 调度器适合于桌面多任务及媒体应用。