

Cisco Catalyst 3750 系列交换机可以为大学、大型企业和运营商提供速率限制

概述

速率限制是大学、大型企业、运营商等组织的 IT 管理员的重要工具。对特殊的用户和端口限制带宽有助于控制网络拥塞，确保很高的性能，创建有效的网络，以及防止少量用户独占网络带宽。在推行千兆位到桌面 (GTTD) 计划的过程中，越来越多的组织开始在园区中安装千兆位以太网端口，因而速率限制变得更加重要。IT 经理将需要管理这些千兆位以太网端口所占用的带宽，直到网络的分发和核心部分通过升级，可以支持更高的带宽需求为止。Cisco Catalyst 3750 系列中的所有交换机都可以支持一套范围广泛的速率限制功能集，它可以用于千兆位以太网和快速以太网接口。

大学生是一种对网络资源要求很高的网络用户。有些学生在他们的宿舍中建立 Web 和 FTP 服务器。有些学生则从互联网上下载很大的文件，例如电影和音乐文件。还有一些学生利用他们的 PC 摄像机拍摄视频。随着他们将这很大的媒体文件与校园内外的学生、家人和朋友共享，流量将会迅速增加。如果不进行速率限制，即使只有少数学生下载和发送很大的文件，都会导致网络速度的急剧下降。当网络在工作时间速度变慢时，教

师可能无法下载重要的研究数据，因此这个问题会变得更加严重。

运营商的宽带接入网络遵从同理。

带宽管理不仅对于大学非常重要，对于企业环境同样也很重要。从互联网上下载大型文件的员工可能会独占一个带宽很小的、通往互联网或者企业内联网的 WAN 连接。采用速度较高的交换机和 PC 网络接口卡 (NIC) 的部门可能会控制企业骨干网和 WAN 连接，而其他部门则会遭受损失。

由于网络带宽不足而导致的网络拥塞可能会导致实时应用（例如 IP 电话和视频）的质量的大幅降低。与数据流量不同，这些实时应用对于由于网络拥塞导致的延时和抖动的承受能力非常低。如果不进行速率限制，从网络服务器上下载大型文件的用户可能会严重地影响他们的相邻用户的电话交谈或者视频会议。较低的网络可用性和性能可能会带来大量的 IT 投诉、降低员工的生产率，导致网络用户满意度的降低。

速率限制

利用速率限制，IT 经理可以控制输入和输出流量的速率，以确保用户或者应用不会超过所分配的最大传输速率，或者独占网络带宽。

这种控制适用于快速以太网,更加重要的是,它适用于带宽更高的千兆位以太网端口。IT经理可以指定策略,为特定的用户、用户群组或者应用分配或高或低的带宽。例如,某个IT经理可能决定只向千兆位以太网端口上连接的教师用户分配200Mbps带宽。结合其他功能(例如基于时间的访问控制列表(ACL)),速率限制可以帮助IT经理精确地管理用户使用带宽的数量和时间,从而可以通过推迟购买更多WAN连接而节约资金。速率限制是一种强大的工具,可以帮助IT经理控制网络,在网络故障发生之前采取有效的预防措施。

Cisco Catalyst 3750系列交换机的速率限制的主要构成包括:

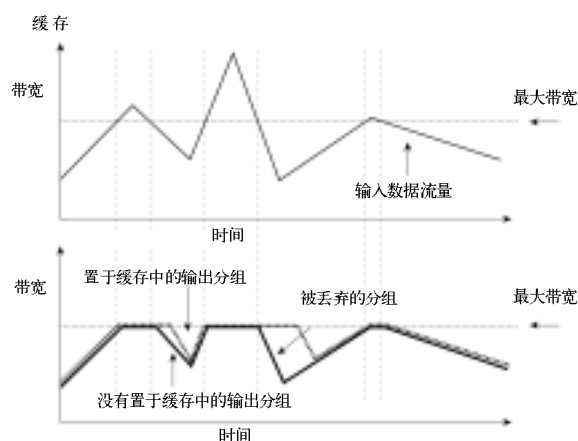
- 缓存
- 队列阈值
- 队列调度
- 流量整形
- 流量监管

所有这些构成要素结合在一起,可以为大学和大型企业用户提供最佳的速率限制解决方案。

缓存

足够大的缓存对于在一个严重阻塞的网络中最大限度降低被丢弃的分组的数量非常重要。(如图1所示)通过将流量放在缓存中排序,较小的突发性流量不大可能导致重发。如果没有缓存,突发性流量可能会导致重发,从而进一步加剧网络拥塞。

图1



在图1中,上面的图形显示的是输入数据流量随时间的变化趋势。有些输入的数据流量超过了允许的最大带宽。在下面的图形中,加粗的黑线显示的是没有置于缓存中的输出分组。如果不进行速率限制,超过最大可用带宽的分组就会被丢弃。点线显示的是置于缓存中的输出分组。利用速率限制,超过最大可用带宽的分组将会被放在缓存中,稍后发送——不会丢失任何一个人组,而且放在缓存中的输出流量也更加平稳。客户可以有选择地为那些拥有更多的突发性流量,或者对抖动和延时更加敏感的队列和端口设置较大的缓存。通常,通往服务器的队列或者专门用于流媒体的队列需要较大的缓存。

队列阈值

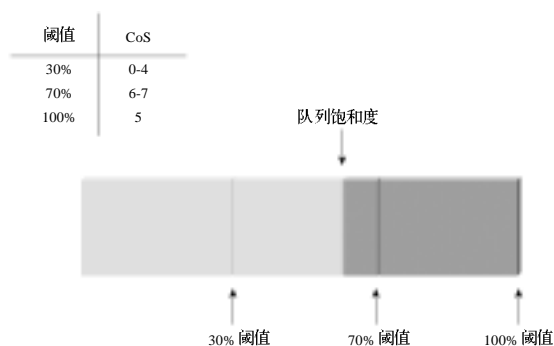
Cisco Catalyst 3750系列交换机支持加权队尾丢弃(WTD)和三个阈值。其中的两个阈值可以由用户设置;第三个设为100%。超过这些指定阈值的分组将会被丢弃。

分组排序的依据是它们的服务质量(QoS)标签。例如,在下面的图2中,服务等级(CoS)为0到4的分组被分配一个30%的阈值,CoS值为5的分组被分配个100%的阈值,而CoS值为6或7的分组则被分配一个70%的阈值。



当分组达到队列的 30% 时，CoS 值为 0 到 4 的分组将会被丢弃。当分组达到队列的 70% 时，CoS 值为 6 或 7 的分组将会被丢弃。只有在队列全满的情况下，CoS 值为 5 的分组才会被丢弃。通常，语音分组的 CoS 值为 5。控制流量和视频流量所分配的 CoS 值为 6 或 7，而低优先级数据流量的 CoS 值为 0 到 4。

图 2 队列阈值



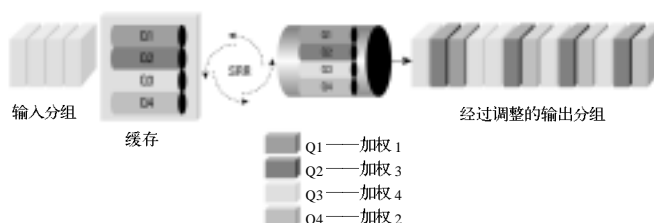
当队列饱和度超过 30% 时，CoS 值为 0 到 4 的分组会被丢弃
 当队列饱和度超过 70% 时，CoS 值为 6、7 的分组会被丢弃
 当队列完全饱和时，CoS 值为 5 的分组会被丢弃

整形循环队列调度

整形循环（SRR）队列调度可以实现更加平稳的流量，因为各个队列将以循环的方式，按照指定的队列权重获得服务。没有一个队列会独占带宽，队列权重较高的队列将传输更多的分组。

Cisco Catalyst 3750 系列的队列调度是利用 SRR 算法进行的。（如图 3 所示）每个队列会被分配一个加权值（缺省情况下四个输出队列的加权值相等）。

图 3 SRR 队列调度



在图三这个例子中，Q1 分配的队列加权为 1，Q2 分配的队列加权为 3，Q3 分配的队列加权为 4，而 Q4 分配的队列加权为 2。队列以循环方式获得服务：Q1、Q2、Q3 和 Q4。在第二个循环中，只有 Q2、Q3 和 Q4 可以获得服务，因为只有 Q1 的加权值为 1。在第三个循环中，只有 Q3 和 Q4 可以获得服务，因为 Q1 的加权值为 1，Q4 的加权值为 2。在第四个循环中，只有 Q3 可以获得服务，因为它拥有最高的加权值。在下一个循环中，这个过程将重复进行。

流量整形

流量整形可以降低网络拥塞，将流量传输变得更加平稳，从而提高带宽的利用率。流量整形将把分组放在一个队列中，并在队列头部安放一个整形标志，以稳定和调整进入网络的流量的速率和数量。对于任何超过指定的传输速率的流量，传输队列将不会传输任何排序的流量数据，直到数据传输速率符合指定的速率。流量整形可以消除突发流量的影响，保证网络中数据流量的稳定性。

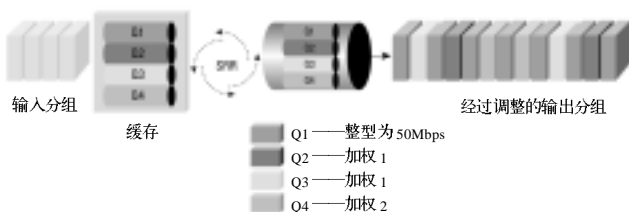
Cisco Catalyst 3750 系列可以利用基于硬件的 SRR 队列调度，在所有四个输出队列上支持输出流量整形。排序和缓存对于输出端口特别重要，尤其是对于速度不匹配和多对一的拥塞。如果某个千兆位服务器向一个 10/100 端口发送一些突发性流量，一些流量将需要存放在缓存中。

Cisco Catalyst 3750 系列支持集中和单独的速率整形。利用集中速率整形，用户可以指定一个端口上的所有四

个队列的最大总传输速率。例如，如果端口 5 的最高带宽是 80Mbps，那么端口 5 上的 Q1、Q2、Q3 和 Q4 的总带宽就不能超过 80Mbps。对于单独的速率整形，用户可以指定每个队列的带宽。例如，端口 4 的 Q1 可以分配 20Mbps，端口 7 的 Q2 可以分配 5Mbps。

用户最多可以对四个输出队列进行整形。没有进行整形的队列可以由 SRR 共享和调度。例如，在下面的图 4 中，Q1 被整形为 50Mbps。这意味着 50% 的输出分组都来自于 Q1。Q2、Q3 和 Q4 则按照它们的队列加权，按照循环的方式共享剩下的 50Mbps。

图 4 流量整形



流量监管

流量监管类似于流量整形。流量监管可以控制进入网络的流量的速率和数量，并由交换机进行处理。但是，流量监管不会使用缓存——超过最大传输速率的分组传输将会被丢弃。

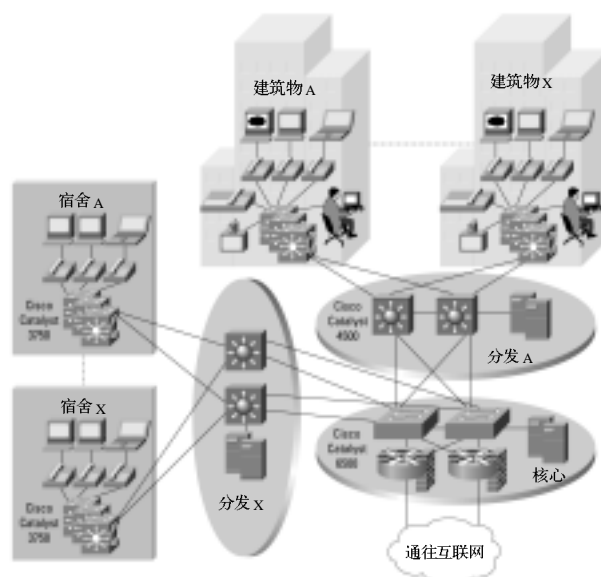
Cisco Catalyst 3750 系列可以利用基于硬件的 SRR 支持输入监管，确保稳定、可预测的转发性能。在单个交换机中，不需要在输入端口进行流量整形和缓存，因为 Cisco Catalyst 3750 交换矩阵是无阻塞的。从一个用户或者服务器端口发来的分组可以立即进入高速矩阵。

每台 Cisco Catalyst 3750 交换机拥有两个输入队列。一个输入队列专门用于严格优先级排序。利用严格优先级排序，这个队列中的所有流量都将优先传输。所有高优先级流量（例如语音流量）都应当放在这个队列中。所有其他流量则放在另外一个队列中传输。

针对大学的速率限制

速率限制在大学中非常重要。利用流量整形和监管，IT 经理可以防止少数使用量很大的用户独占网络带宽。此外，目前有很多大学目前正在试用 GTTD，而速率限制可以为 IT 经理提供一个重要的工具，帮助他们管理这些千兆位以太网端口所占用的带宽，直到网络的分发和核心部分升级到更高带宽的技术（例如 10Gb 以太网）为止。教师和学生发送和接收的数据不能超过他们所分配到的带宽。IT 经理可以选择为教师和研究学生提供比本科生更多的带宽，并可以为应用服务器和 workstation 分配更多的带宽。

图 5 大学网络



宿舍

在宿舍，IT 经理可以限制每个学生所获得的带宽，以确保没有学生独占网络带宽。一个在宿舍中建立音乐共享服务器的学生不仅会降低整个宿舍网络的速度，还会影响整个校园网络。数以千计的想要从某个宿舍服务器



上载或者下载文件的学生将会严重地影响分发网络和核心网络的上行流量。这可能会占用大部分 WAN 连接，这些连接通常只是一些 T1/E1 或者 T3/E3 线路。

为了防止发生这种情况，IT 经理可以将每个学生所获得的带宽限制到一个普通学生的级别。速率限制可能是不对称的。普通学生使用的下行流量多于上行流量，因此 IT 经理可以为每个学生设置较多的下行流量和较少的上行流量。IT 经理可以利用介质访问控制 (MAC) 地址跟踪用户和计算机，并可以对用户或者计算机所获得的带宽数量进行速率限制。

视频是最占用带宽、对抖动最敏感的应用之一（例如学生通过下载课堂视频，在他的宿舍房间中上课）。当 IT 经理对下行和上行带宽进行速率限制时，他们必须为那些需要观看视频（通常介于 300Kbps 到 2Mbps 之间，具体取决于所采用的加密方式）的学生提供足够的下行流量。在这种情况下，IT 经理应当为视频队列分配更多的缓存，并将视频流量分配到高优先级队列，从而最大限度地减少抖动。

校园建筑物

在校园建筑物中，IT 经理通常需要对教师流量分配较高的速率，并为教师流量设置一个高于学生流量的优先级。千兆位以太网端口上的教师流量可能会限制于 200Mbps，并且分配到次高优先级队列。在学生的教室中（讲课地点），连接到同一个千兆位以太网端口的学生的带宽可能会被限制为 500Kbps。但是，在研究生的研究中心和学生工作站集群中，学生获得的带宽可能会高一些，例如 4Mbps。工科研究生通常需要发送很大的设计文件到服务器进行处理，学习制图的学生需要传输很大的视频文件，而工作站集群中的本科生需要完成他们的课外作业和项目工作。

在宿舍和校园建筑物中，IP 电话流量都应当进行速率整形，并放在严格优先级队列中。严格优先级队列确保了语音流量可以获得最高的优先级。严格优先级队列可以为语音流量提供专用的带宽，从而最大限度地降低延时和抖动。尽管语音流量对延时和抖动非常敏感，但是它并不需要很高的带宽（低于 100Kbps）。在图 6 显示的这个例子中，语音流量被整形到 25Mbps。经过整形的带宽是用户获得理想的语音流量传输的重要因素。

图 6 大学网络中的速率限制



IT 经理可以为一个队列采用不同的丢弃阈值。例如，对于图 6 中的 Q2，IT 经理可以为视频流量分配 100% 的丢弃阈值，为教师 HTTP 流量分配 40% 的丢弃阈值。视频流量是一种实时应用，对于丢弃分组非常敏感。另一方面，HTTP 流量通常由较短的突发性流量构成，对于丢弃分组的敏感度较低。为 HTTP 分组重传等待一到两秒钟对用户体验的影响远远比不上为视频分组重传等待一到两秒钟的影响大。

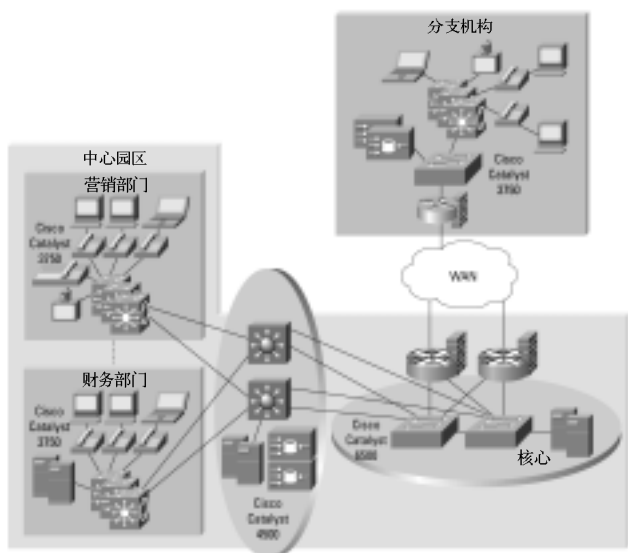
IT 经理可以优先为指定的建筑物或者分发中心提供较高的带宽，具体取决于带宽需求，或者支付更多资金的系可以获得较高的速率限制。一般而言，IT 经理为校园建筑物分发网络与核心网络间流量分发的带宽要高于宿舍分发网络与核心网络之间的流量。教师办公室、研究生研究中心和学生活动中心都位于校园建筑物中。这些使用量很大的中心通常需要较高的网络带宽，并且对通往互

联网的WAN连接的需要也较高。速率限制让师生们可以更加有效地使用网络带宽。

针对企业的速率限制

与大学环境一样，速率限制也可以用于在企业中进行智能化地带宽分配。他可以防止个别部门或者员工控制全部的网络带宽——它让IT经理可以为需要更多带宽的部门和应用分配更高的带宽。

图7 企业网络



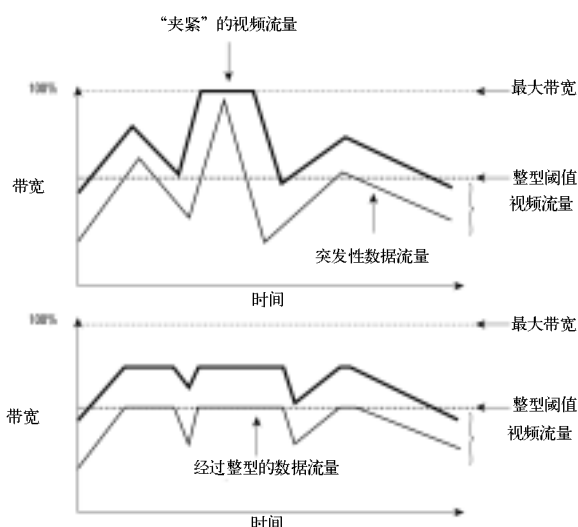
中心园区

在很多组织中，有些部门对于Web或者内联网接入的需求高于其他部门。营销部门经常需要在互联网上搜索竞争产品的数据和行业报告。档案部门可能需要从位于另外一个城市的企业数据库下载数据。IT经理可以从分发网络对营销部门进行速率限制，使其可以获得较高的HTTP流量带宽，并可以对档案部门进行速率限制，提高它与某个特定服务器之间传输内联网流量的带宽。

一个使用量很大的网络用户不应当影响其他员工的网络质量。例如，工程师们经常要执行批量任务。尽管批量任务有助于提高工程师的工作效率，但是某个利用文件传输协议（FTP），从工程数据库服务器发送20个1GB的设计文件的工程师不应当影响另外一个以500Kbps的速度观看培训视频的员。文件下载的速率必须受到限制，因为像培训视频这样的实时应用对于由突发性流量导致的抖动非常敏感。为了解决这个问题，IT经理们可以将视频流量分配到一个优先级较高的队列，并为视频流量队列分配较大的缓存。流量整形有助于减轻由于大型文件传输所导致的突发性流量所带来的影响。（如图8所示）



图8 流量整形提高网络流量的平稳度



在图8中，数据流量没有经过整形，并且存放在一个优先级高于视频流量的队列中。对于保证良好的用户体验，视频流量需要稳定的带宽。每几秒钟就暂停一次的视频将会非常影响观众的情绪。在图8的上图中，存在一个很大的突发性数据流量（视频流量在最大带宽处被“夹紧”）。当所需要的带宽超过可用带宽的最大值时，视频分组会被丢弃，员工将会收看到抖动的视频和不完整的视频传输。在下面的图形中，数据流量得到了整形。超过整形阈值的流量会被存放在缓存中，稍后传输。数据流量的传输速率比上图平稳得多。没有视频分组被丢弃，因此视频体验不会受到影响。

分支机构

分支机构的资源更加有限。尽管有些工作可以独立完成，但是分支机构的员工经常需要从中央园区收发数据。根据工作性质的不同，员工可能需要通过内联网，向总部发送一些很大的文件或者很多小文件。由于只有一条T1（速率为1.5Mbps）连接通往总部，一个1GB的文件

在全速情况下需要1.5小时才能传完（这还必需要求在这1.5个小时内没有其他人使用内联网）。这可能会极大地降低分支机构的员工的生产率。

IT经理和员工可以通过购买更多的WAN带宽，更好地管理带宽分配和文件传输时间，解决这个问题。购买两条T1线路可以将文件传输时间缩短一半。IT经理还可以对员工进行速率限制，确保任何一个员工都不能独占WAN连接。例如，没有一个员工可以一次使用超过250Kbps的WAN流量。IT经理可以在不同的时段向员工分配不同的速率限制。IT经理可以在白天分配低于夜晚的速率限制。员工也可以将他们的文件传输任务安排在夜间进行，即在其他员工回家之后。

通过智能化地控制来自于用户和应用的流量，速率限制可以降低一个组织的整体运营成本。它让企业客户可以在不提高总上行带宽的情况下，移植到千兆位以太网。通过让IT经理无需立即移植到10GB以太网上行链路，就可以部署千兆位以太网交换机，企业可以在节约资金的同时，最大限度地提高网络的可用性和性能。

总结

速率限制是一个非常重要的工具，可以帮助IT经理加强网络管理，降低流量拥塞——在大学和企业环境中提高带宽的利用率。速率限制将控制权交给了IT经理，让他们可以主动地管理网络，而不是被动地在某个由于网络拥塞导致地问题发生之后设法解决问题。速率限制还可以通过推迟企业对于更多WAN带宽的需求而降低成本。

思科在你身边 世界由此改变



思科系统(中国)网络技术有限公司

北京

北京市东城区东长安街一号
东方广场东方经贸城
东一办公楼 19~21 层
邮政编码:100738
电话:(8610)65267777
传真:(8610)85181881

广州

广州市天河北路233号
中信广场43楼
邮政编码:510620
电话:(8620)87007000
传真:(8620)38770077

上海

上海市淮海中路222号
力宝广场32~33层
邮政编码:200021
电话:(8621)33104777
传真:(8621)53966750

成都

成都市顺城大街308号
冠城广场23层
邮政编码:610017
电话:(8628)86758000
传真:(8628)86528999

如需了解思科公司的更多信息, 请浏览 <http://www.cisco.com>

2003年思科系统(中国)网络技术有限公司北京印刷, 版权所有。

2003© 思科系统公司版权所有。该版权和/或其它所有权利均由思科系统公司拥有并保留。Cisco, Cisco IOS, Cisco IOS 标识, Cisco Systems, Cisco Systems 标识, Cisco Systems Cisco Press 标识等均为思科系统公司或其在美国和其他国家的附属机构的注册商标。这份文档中所提到的所有其它品牌、名称或商标均为其各自所有人的财产。合作伙伴一词的使用并不意味着在思科和任何其他公司之间存在合伙经营的关系。