**TC原理及在云平台中的应用**

1. **TC技术背景**

TC（traffic control）是Linux中用于进行流量管理的工具，工作在数据链路层和物理层之间，可以对流出网卡的流量进行限速、应用策略等，为Linux提供QoS服务。

报文在交付给网卡之前，会先进入内核中的TC模块，流量控制的一个基本概念是队列(Qdisc)，每个网卡都与一个队列(Qdisc)相联系，每当内核需要将报文分组从网卡发送出去，都会首先将该报文分组添加到该网卡所配置的队列中，由该队列决定报文分组的发送顺序。因此可以说，所有的流量控制都发生在队列中。对于队列的各项流控功能，需要通过类别（Class）和过滤器（Filter）来实现。过滤器是队列用来对数据报文进行分类的工具，它决定一个数据报文将被分配到哪个类别中；而类别定义了具体的流控规则，比如限速。

常用于限速功能的队列是HTB队列，该队列是一个可分类的队列，使用令牌桶规则。以下通过实例来进行用法解释。

* 为网卡eth0添加HTB队列

tc qdisc add dev eth0 root handle 1: htb default 1

其中root表示该队列是一个根队列，handle 1: 表示队列的句柄为1，htb表示这是一个HTB队列，default 1表示未分类的流量分配给类别1:1

* 为根队列创建类别，限速100M比特每秒

tc class add dev eth0 parent 1:0 classid 1:1 htb rate 100Mbit ceil 100Mbit

其中parent 1:0表示该类别的父队列为根队列1:0，classid 1:1表示创建一个标识为1:1的类别，rate 100Mbit表示确保该类的带宽有40M比特每秒，ceil 100Mbit表示该类的最高可占用带宽为40M比特每秒

* 为类别设置过滤器

tc filter add dev eth0 parent 1:0 protocol ip prio 1 u32 match ip dst 0.0.0.0/0 flowid 1:1

protocol ip表示该过滤器应该检查报文分组的协议字段，prio 1表示过滤器优先级，对于不同优先级的过滤器，系统将按照从小到大的优先级顺序来执行过滤器，u32及后边内容是选择器，用来匹配不同的数据流，match ip dst 0.0.0.0/0意味着匹配所有发出的数据流，flowid 1:1意思是将匹配的数据流分配给类别1:1

通过队列、类别、过滤器的配置，可以实现对主机的带宽限制。

1. **使用TC对云主机进行限速**

计算节点典型的网络架构如下图所示



云主机的网卡连接了宿主机的一个tap设备，所有云主机的流量都通过该tap流入流出，如果在tap设备上使用tc规则，则可以控制流经该设备的数据流，达到对云主机限速的目的。