

MPLS 和 Internet 流量工程

MPLS & Internet Traffic Engineering

孟 魁 张根度

(复旦大学计算机科学系 上海 200433)

Abstract Rapid growth of Internet traffic and increasing requirement for service quality have made traffic engineering an essential consideration in the operation of large Internet backbone networks. Historically, traffic engineering in the Internet has been hampered by the conventional IP technologies. This article discusses the challenge of Internet Traffic Engineering and the applications of MPLS to traffic engineering in IP networks.

Keywords Traffic engineering, Internet, MPLS

随着经济全球化趋势的进一步演进, Internet 产品和商业工具的不断更新以及新经济对社会生活的介入, Internet 用户的数量有了极大的增长, 整个网络的主干链路上每个用户的平均数据流量以指数规律上升, 与此同时, 新的网络应用和网络业务的层出不穷, 更加剧了网络发展的日趋大型化、复杂化。

当前 ISP 面临的问题是如何满足客户对高质量服务和多样化服务的需求, 以及如何最有效地利用高成本的网络资源。这就要求 ISP 能够对网络中的客户业务流进行明确的控制, 根据客户的业务性质、要求的业务级别提供相对应的特殊路径。同时要能在网络中不同的链路、路由器和交换机等资源之间平衡业务负荷, 使网络各主干线路上的业务量均匀分布, 达到一定负荷(如 60%), 且不会出现拥塞。ISP 还希望能采取一种灵活的、可扩展的方法进行网络管理, 以适应将来网络发展的需求。

可是, 传统 IP 技术面向无连接的无序性网络管理对此是无能为力的, 而流量工程技术使上述想法的实现成为可能。过去, Internet 流量工程的发展受到传统 IP 技术局限性的阻碍。而近年来, 多协议标签交换(MPLS, Multi-protocol Label Switching)以其各方面的优势脱颖而出, 成为 IP 网络上强有力的流量工程解决方案。

1 流量工程

将业务流映射到现有物理拓扑上的任务被称作流量工程。流量工程的目的是提高网络运行的有效性和可靠性, 同时优化对有限的网络资源的利用, 提高网络

资源利用率。流量工程试图让实际网络业务量以一种最优的方式存在于物理网络之中, 对 ISP 有相当重要的意义。

流量工程实际上是一套工具和方法, 无论是在网络设备和传输线路正常还是失效的情况下, 它都能从给定的基础设施中提取最佳的服务。也就是说它要对已安装的资源进行优化, 事实上, 它是对网络工程或网络规划的一种补充和完善措施。

流量工程有传输业务和资源利用两方面的目标。前者的度量包括有关业务流 QoS 的指标, 如: 最小丢包率, 最小时延, 最大吞吐量和服务等级协议的执行情况等。能够在同样业务量的情况下, 提高 best effort 型业务的传输质量, 而不是单纯通过牺牲其它业务而提高某些有特殊需求的业务的服务质量, 是流量工程的一个主要目标, 也是检验网络整体性能的一个重要指标。后者则要求对网络资源进行有效管理, 尤其是要确保网络的有些通路上不出现拥塞或过载, 而其它同样可行的通路上却还未被充分使用的情况。当前网络上, 带宽是重要和稀缺的资源, 因而流量工程的一个重要功能就是要有效地管理带宽。

流量工程包括性能监视系统和网络管理配置系统两个子模块。要想实现业务量的合理分布, 首先必须知道业务量是如何分布的, 也就是说, 需要在每个网络分枝点对网络上实时传递的业务流进行统计, 同时还需要统计网络资源的使用状况。这个统计功能模块是流量工程的一个重要组成, 称为性能监视系统。有了实时获取的网络状态信息, 根据一些预先设置的调整策略(由流量工程师负责制定), 流量工程通过它的另一个

孟 魁 硕士生, 研究方面为计算机网络及应用。

重要组件——网络配置管理系统对网络进行相应调整,使网络达到与流量控制策略相一致的期望的状态。可调整的内容包括:业务量管理参数;与路由选择相关的参数;与网络资源相关的属性和限制。这三个方面概括了与流量工程有关的网络参数,其具体内容与网络所采用的技术有很大关系。通过对这些参数的调整,可以将业务流重新导向、分流,从而减轻网络的局部压力,同时也可以使业务流的传输性能得到改善。

2 Internet 流量工程

流量工程在传统电信商的网络中得到比较广泛的应用,如传统的 PSTN 网和 ATM 网络。由于传统电信网是面向连接的,因此无论是流量测量、预测和控制都是比较容易的。

而 Internet 本质上是无连接的尽力传送网络,要在这样的基础上实现流量工程无疑要面对更多的挑战。首先,Internet 的发展是彼此独立和不连续的,而且 Internet 业务流表现出极大的动态性,分布不均匀,难以进行精确的预测。其次,网络运行环境也处于不断变化之中。几乎每时每刻都有新的网络资源加入或退出,同时新的网络应用更是层出不穷,这也给网络带宽带来了新的压力。同时,由于设备自身的局限或其它原因,一些网络设备并不能位于网络中的最佳位置。而各自治系统之间的业务流穿越更增加了网络的复杂性。

在早期的传统路由器骨干网中,流量控制是通过在网络链路上的量度值进行管理来实现的,这个量度值是由网络管理人员在综合考虑网络的业务分配状况及阻塞发生的可能性后配置的。当链路发生阻塞时,网络管理人员通过更改量度值来解决出现的阻塞。这种方法在中小型网络中可以实现对流量的控制,但在大型网络中则显得力不从心。

事实上,在流量控制中,总是希望人的干预最小化,也就是说至少在一定的情况下,网络能动态地完成对度量值的调整;而且更希望在拥塞发生前就发现,并进行有效的调整。这就要求对网络上的业务流量和网络资源利用情况进行准确、实时的监测和控制。

尽管基于内部网关协议 IGP(如 OSPF、IS-IS)能进行简单的负载均衡(Load Balancing),但是由于没有面向连接能力(不提供显式路由能力),因此利用内部网关协议进行流量工程是比较初级和难以控制的。

传统的路由协议往往只考虑了网络的静态特性(如链路的物理带宽)和网络拓扑的变化(链路的通断状态),没有考虑更详细的网络动态特性,如链路的实际可用带宽、链路传输的动态延迟等。这样带来的弊端是显而易见的。

例如,在下图所示的这个流量工程例子中,有两条

从路由器 A 到路由器 E 的路径:A-B-E 和 A-C-D-E。如果一个路由器选择了其中一条从 A 到 E 的最短路径(A-B-E),那么它就会传送所有预定前往那些通过 E 可以抵达的网络通信,因而在该路径上带来的流量可能会造成拥塞,而另一条路径(A-C-D-E)则负载不足。为了最大限度地提高整个网络的性能,将一部分流量从一条链路转移到另一条链路也许是可取的办法。

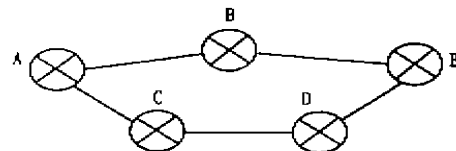


图 1 流量工程举例

虽然人们在这个简单的例子中可以设定 A-B-E 路径的成本等于 A-C-D-E 路径的成本,但这种负载均衡方法在复杂的拓扑网络中会变得十分麻烦(虽然这种方法并非不可能)。在 IETF 提出的多协议标签交换(MPLS, Multi-protocol Label Switching)中,则可以通过使用显式路由(Explicit Routing)简单灵活地解决这一问题。

3 MPLS—多协议标签交换

MPLS(MPLS, Multi-protocol Label Switching)源于早期的 IP 交换解决方案,是一种在开放的通信网上利用标签引导数据高速、高效传输的新技术。

3.1 MPLS 体系结构及有关概念

MPLS 体系结构如图 2 所示。有关术语概念介绍如下:

- * 标签交换路由器(LSR):负责第三层分组转发和第二层分组标签交换的设备。完成 FEC 到 LSP 的建路和下一跳改变的发起。

- * 边界标签交换路由器(LEP):负责从/向 MPLS 域转发分组,并与内部 LSR 通信以交换 FEC/标签绑定信息。完成 FEC 划分,流量工程,LSP 建路发起,IP 包转发、Diff-Serv 等任务,在中间节点的下一跳改变时也为 LSR 提供 FEC 生成及路由重选。

- * 转发等价类(FEC):以相同方法处理并转发的一组 IP 包。在 MPLS 网络中,和标签一一对应。

- * 标签(Label):一个包含在每个分组中的短固定长度的数值、用于通过网络分类转发分组。一个标签只在一对正通信的 LSR 之间起作用,表示属于一个从上游 LSR 流向下游 LSR 的特殊 FEC 分组。这一点对于 MPLS 的可扩展性非常关键。

- * 标签交换通道(LSP):一个从入口 LER 到出口 LER 的交换路径,由 MPLS 节点建立,以标签交换的

方式转发一个特定的 IP 包(FEC 分组)。

* 标签分发协议(LDP):用于在 LSR 之间分配和分发 FEC/标签绑定信息的控制协议。一旦 LDP 完成它的任务,从入口到出口就一条 LSP。

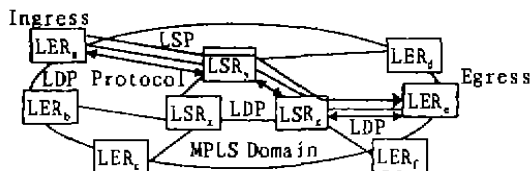


图 2 MPLS 体系结构

3.2 MPLS 工作原理

MPLS 是一种交换和路由的综合体,它将第三层(路由层)的智能、灵活性和可扩展性与第二层的交换机制(不包括它面向连接的服务)结合起来。一个 MPLS 网络采用标准分组处理方式对第三层的分组进行转发,采用标签交换对第二层分组进行交换。它基于对等模型,即每一个 MPLS 设备运行一个单一的 IP 选路协议、交换选路表更新信息,并维护一个拓扑结构和一个地址空间。

当分组进入 MPLS 网络时,入口 LER 检查分组的多个字段来确定分组属于哪一个 FEC。如果存在一个 FEC/标签关联,入口 LER 就给分组贴上一个标签然后引导到合适的出口。LSR 利用标签索引一个出口端口和一个新标签的连接表,进行标签交换后,把分组从出口端口转发到下一跳路由器。当分组到达出口 LER 时,标签被去除,并进行第三层分组的处理。

4 MPLS 在 Internet 流量工程上的优势

MPLS 最有用的应用之一就是流量工程。

不同于传统的基于目的地址的最短路径路由模式的是,MPLS 允许在入口 LER 处根据业务流的特性决定是否建立显式 LSP(人工预定或利用选路协议)。通过显式 LSPs,就使无连接的 IP 网络有了类似于电路交换的能力。由于支持显式路由,MPLS 增强了 IP 网络流控和自愈恢复能力。

除支持显式路由外,MPLS 还具有下列适合于流量工程、尤其是大型网络和 Internet 的流量工程的特性:

1)MPLS 提高了网络的可监测性。MPLS 在一个无连接的网络中引入了连接模式的特性,且基于拓扑驱动,使网络性能稳定,有效地解决了 N² 问题。

2)更好地支持 QoS,能提供更多高质量的网络服务。由于支持显式路由,MPLS 允许个别用户流获得指定级别的服务,MPLS 首先在网络边缘识别客户流(通过分组头中的多个字段),决定是否建立显式路径,然

后将相应的标签附加到分组头,通过 LSP 发送出去。

3)导入和实施新型网络业务将会有更大的灵活性。MPLS 不但支持多种网络层技术,而且是一种与链路层无关的技术,它同时支持 X.25、帧中继、ATM、PPP、SDH、DWDM……,保证了多种网络的互连互通,使得各种不同的网络传输技术统一在同一个 MPLS 平台上,不仅支持各种类型的传输数据,也便于各种新业务在基于 MPLS 和 IP 网上实现。

4)提高了网络扩展性,适合构建大型网。MPLS 提供多种路由协议,支持大规模层次化的网络拓扑结构;同时 MPLS 进行了很好的网络功能划分(核心 LSR 和边缘 LER),使网络具有良好的扩展性,满足大规模网络的各种要求(如灵活性、可扩充性与可管理性等要求)。

应当说,MPLS 是在 IP 网络中提供流量工程的重要工具。虽然在 MPLS 之前的综合服务(Integrated Services)和差异性服务(Differentiated Services)能够解决一部分服务质量的问题,但只有 MPLS 才是一种最全面的服务质量保证体系。通过 MPLS 的流量工程技术,传统的 Best Effort IP 网络可以平滑地演进为受控的多业务网络。MPLS 被认为是有效解决流量工程的一个理想方案。

5 MPLS 流量工程模式

在介绍 MPLS 流量工程模式前,先介绍几个术语。

Traffic Trunk 是流量工程中一个非常重要的概念,是一个同类业务流的“集合”。流量工作的中心任务就是把 Traffic Trunk 映射到网络拓扑上去。在 MPLS 网络中,是通过显式 LSPs 的路由选择来将 Traffic Trunk 映射到网络拓扑上的。在 MPLS 中,这两者的结合被称为 LSP 隧道和 Traffic Trunk 隧道。

LSP 隧道允许用多种方法对网络性能进行优化。例如,如果因不合适的路由导致的拥塞被发现的话,LSP 隧道可以被重新路由来避免拥塞。在两个结点间还可以建立多个 LSP 隧道,根据本地策略将两结点间的业务分配到这些隧道中。

一个 MPLS 流量工程模式包括以下四个基本的功能模块:

1)路径管理模块 涉及显式路由选择,LSP 隧道建立和维护。一个路径管理策略不仅定义了路径选择标准,还定义了维护已建 LSP 隧道的准则。路径管理模块有三个基本功能:路径选择、路径建立、路径维护。

一些有关 LSP 隧道和资源的属性与该模块有关。其中 LSP 隧道属性包括:

• 流量参数(Traffic Parameter):用于表征 LSP

隧道的流量特性,如峰值速率、平均速率、允许的突发速率或 LSP 隧道带宽等,可据此计算带宽需求,有效分配带宽并避免拥塞。

- 适应性 (Adaptivity): 表示 LSP 隧道是否能随网络状况的变化进行自主优化。一个适应的 LSP 隧道在出现更好的路由时可以自动地重新选路。而非适应性 LSP 隧道一旦确定路径后就不会改变。

- 优先级属性 (Priority): 当多个 LSP 隧道抢占资源,决定哪条 LSP 隧道将获得资源。

- 资源强占性属性 (Preemption): 在资源不足的情况下规定该 LSP 隧道是否可能强占其它 LSP 隧道的资源或是否允许其它 LSP 隧道强占自己的资源。对于网络而言,强占并不有利于网络资源的优化,但它可以满足部分应用的特殊需求。

- 自愈性属性 (Resilience): 表明 LSP 隧道对其路由被破坏时是否重新选路。

- 资源亲和级别属性 (Resource Class Affinity): 根据 LSP 隧道与资源的关系密切程度,决定在该 LSP 隧道进行路径选择时是否必须包含或不可包含该资源。

资源属性包括:最大分配系数 (Maximum Allocation Multiplier, MAM),缺省流量工程度量值 (Default Traffic Engineering Metric) 以及资源级别 (Resource Class) 等属性。

2) 业务流分配模块 一旦一个 LSP 隧道建立,就应当为它分配业务流,业务流分配模块就是负责为建立的 LSP 隧道分配业务。它有分割和分配两大功能:业务流模块根据一定的原则对流入的业务流进行分割,再按相应的业务流分配规则把分割后的业务分配到已建的 LSP 隧道上。

解决业务流自动分配的一种方法是把 LSP 隧道看成是 IGP 域内的直通路径。当存在多条路径通往某一给定结点时,可使用筛选规则等一些附加属性控制业务流的分配。在两个结点之间的多个 LSP 隧道中进行负载分配是业务流分配的一项重要工作。这可以通过显式或隐式地指定每一个 LSP 隧道的权值,根据权值进行业务流分配或者根据网络状态反馈调整业务流分配。

3) 网络状态信息分发模块 负责 MPLS 域中拓扑状态信息的发送。它利用在扩展的传统 IGP 进行网络链路状态传播时附加一些网络状态信息完成这项工作。这些附加信息包括:最大连接带宽,最大分配系数 (Maximum Allocation Multiplier, MAM),缺省流量工程度量值 (Default Traffic Engineering Metric),每个优先级的保留带宽以及资源级别 (Resource Class)

等。

拓扑状态信息用于对 LSP 隧道进行限制路由。

4) 网络管理模块 网络管理是 MPLS 进行流量控制的重要工作。MPLS 实施流量工程的优势就在于它便于对网络进行监测和控制。一个 MPLS 网络管理系统一般有下列功能:配置管理功能,性能和计费管理功能,缺省管理功能。

总的来说,这些功能可以得到 MPLS 网络管理对象如 LSP 隧道的状态,并控制它们的特性,从而达到控制网络性能的目的,如:点到点的业务流特性可以通过对 LSP 隧道的监测统计得到;丢包率可以通过对 LSP 隧道两端的出入流量统计差异估算得到;包延迟可以通过向 LSP 隧道发送的探测包传输时间的测定进行估算;当 MPLS 管理对象的状态超过规定阈值时,系统将会发出事件通告等。

要单纯依靠 MPLS 流量工程来实现对大型网络的性能优化是比较复杂和困难的,一般还会添加一些离线流量工程支持工具来提高 MPLS 的在线流量控制能力。这些离线工具与 MPLS 网络管理系统相连接以提供外部反馈控制。

结束语 总而言之,MPLS 在一个无连接的网路中引入了连接模式的特性,减少了网络复杂性,兼容现有各种主流网络技术,有效降低了网络成本,能提供更多高质量的网络服务,同时满足了 ISP 和网络最终用户的需要。

我们可以这样认为:MPLS 以比较小的代价赋予互联网一种面向连接的控制能力,从而使互联网中的流量工程成为可能,而流量工程反过来也成为促进 MPLS 的主要动力之一。

参考文献

- 1 Awduche D, et al. Requirements for Traffic Engineering Over MPLS. RFC 2702, Sept. 1999
- 2 Feldman, et al. Evolution of Multiprotocol Label Switching. IEEE Communications, 1998, 36(5)
- 3 Rosen E, Viswanathan A, Callon R. Multiprotocol Label Switching Architecture. IETF Internet draft, work in progress, July 1998
- 4 Awduche D, et al. MPLS and Traffic Engineering in IP Networks. IEEE Communication Magazine, Dec. 1999
- 5 Swallow G, et al. MPLS Advantages for Traffic Engineering. IEEE Communication Magazine, Dec. 1999
- 6 Metz C T. IP 交换技术协议与体系结构 (IP Switching Protocols and Architectures). 吴靖, 龚向阳, 金跃辉, 张鹏, 等译. 机械工业出版社, 1999