

# IP 网络生存性研究综述<sup>\*</sup>

吴 静 郭成城 晏蒲柳

(武汉大学电子信息学院 武汉 430079)

**摘 要** 随着交互式应用和各种实时业务的增长对网络生存性要求的提高,IP 网络的生存性研究受到了越来越多的关注。本文首先对 IP 网络生存性面临的问题以及影响因素作了概括,并且对目前有关纯 IP 网络生存性、MPLS 网络生存性、IP 网络与底层的生存性协调研究作了归纳总结,尤其对纯 IP 网络的生存性研究现状从多方面进行了比较详细的介绍,最后指出 IP 网络的快速重路由机制具有很好的实际应用潜力,保护恢复能力和服务质量在 MPLS 网络中的结合、动态多层保护则是未来 IP 骨干网生存性维护的发展方向。

**关键词** 生存性,纯 IP 网络,多协议标记交换(MPLS),多层流量工程

## Overview of IP Network Survivability

WU Jing GUO Cheng-Cheng YAN Pi-Liu

(School of Electronic Information, Wuhan University, Wuhan 430079)

**Abstract** Optical network survivability has been a hot topic of research for a long time. However, it is only recently that network survivability has attracted the attention of large research community. The growth of business-critical applications is the primary driver behind this increased interest in network protection and restoration. MPLS network is connection-oriented and its protection and restoration mechanism is easy to be configured. But most IP networks are not updated to support MPLS, so it is meaningful to study the survivability of pure IP network. There are many challenges for IP network survivability which are generalized in the paper above all. And then research on different IP network survivability and the coordination with lower layer survivability are concluded, especially the state of the art for pure IP network is introduced in detail. Finally it is indicated that IP fast rerouting is potentially efficient for implementation, the recovery and QoS considered together in MPLS network is preferred and multi-layer resilience would be the hot topic in the future.

**Keywords** Survivability, Pure IP network, Multi-protocol label switching(MPLS), Multi-layer traffic engineering

## 1 引言

近年来,电子商务、虚拟社区等要求高可靠性的业务和实时业务的增长对网络的保护和恢复提出了更高的要求<sup>[1,2]</sup>。研究表明,当网络可用率为 99% 时,大型电子商务网站一年的损失将达到 370 万美元。要避免巨大损失,电子商务公司需要可用率达到 99.999% 的网络<sup>[3]</sup>。然而,现代网络并没有达到应用业务的需求。文[4]指出,在一个重要的 ISP 中,20% 的链路失效时间小于 1 天,70% 的链路失效时间小于 10 天;因特网骨干网路径的平均失效时间为 2 天,只有大约 20% 的路径在 5 天内没有变化。造成网络故障的原因很多,包括光纤断裂、错误配置、路由器不稳定等等。随着器件可靠性的增强,现在网络的故障往往由于人为的操作维护所引起,例如协议配置不匹配、路由器软件升级失误、隧道配置造成环路等等。G. Iannaccone 等人通过对 Sprint 网络持续 5 个月的观察,发现 45% 的失效事件在维护期间产生<sup>[5]</sup>。N. Dubois 等人对另 1 个大型 ISP 进行连续 1 个月的观察,发现 75% 的 IS-IS 事件是维护操作引起的<sup>[6]</sup>。为此,提高网络的生存能力是保证互联网高效承载新兴业务的关键之一。

网络生存性是为 ITE(Internet Traffic Engineering, 因特

网流量工程)问题推荐使用的功能性标准之一。该标准用于衡量网络在出现故障时保持服务连贯性的能力,要求能从网络故障中迅速恢复,而且要保持当前服务所要求的 QoS (Quality of Service, 服务质量)<sup>[7]</sup>。无论是现在的因特网还是下一代的 NGN 和 NGI,协议结构都呈沙漏装,网络层使用 IP 协议,将异构网络互连起来,而网络层以下和以上有不同的网络技术和协议,因此研究 IP 网络的生存性具有重要的意义。近年来网络生存性研究的对象主要集中在网络层以下,尤其是在光层上,但是随着网络技术和网络应用的不断发展,IP 层网络的生存性研究目前受到了越来越多的关注,这主要归因于以下几个因素:

(1)底层的网络并非全部采用光网络技术,当故障发生在网络层时,就不能依靠光层的恢复机制来解决。

(2)电子商务和各类交互式应用的繁荣和发展对端到端的可靠性保证提出了更高的要求<sup>[8]</sup>。显然,仅靠光层的保护是远远不能满足端到端的要求的。

(3)MPLS 技术为 IP 网络的保护和恢复提供了新的渠道。MPLS 实际上是一种面向连接的网络技术,因此可以通过路径备份,故障时将特定的流切换到特定的路径备份路径中,从而为端到端的网络恢复奠定了基础。RFC3469<sup>[9]</sup> 总结

<sup>\*</sup>国家自然科学基金基础资助(90204008)。吴 静 博士研究生,主要研究方向为网络管理和性能优化、网络路由技术;郭成城 教授,博士生导师,主要研究领域包括网络管理、Web 集群服务器;晏蒲柳 教授,博士生导师,主要研究领域为网络管理。

了各种基于 MPLS 的网络恢复草案,形成了基于 MPLS 网络恢复的整体框架。

(4)发展中的 NGI 和 NGN 都以 IP 技术为核心,IP 网络的生存性关系到下一代网络承载新兴业务的能力。

## 2 影响 IP 网络生存性的因素

影响 IP 网络生存性的因素众多,其中网络拓扑结构和网络协议是两个主要因素。

从图论的角度出发,格形拓扑的网络不易形成单点瓶颈,能提供比较好的生存性。节点度数越高的网络,由于能提供更多的冗余链路,因此保护和恢复的能力也越强。不少学者针对 IP 网络拓扑规划作了深入研究。Minkyu Kim 等人考虑在拓扑上具有随机性的大型网络的保护问题,利用随机图模型刻画这种网络,获取和可靠性相关的几个参数的概率边界<sup>[10]</sup>。他们的研究主要集中在 3 个方面:首先,对于随机选择的节点对,计算节点对之间平均距离的下界,讨论了边界的紧密性;其次,考虑到主要路径和保护路径可能形成环,计算节点对间最短环的下界;最后指出,在给定最大有限长度情况下,被保护的连接非常少。通过网络模型的概化,以上结论也可以适用于度数满足幂律分布的网络。文<sup>[11]</sup>研究如何在不可靠的网络元素基础之上建立高可靠性的物理网络拓扑。该网络模型假设节点不会失效,而链路可能失效,并且考察非独立性和独立性链路失效。文中指出平均链路失效概率小的网络,其可靠性对于网络拓扑结构不敏感,在平均链路失效概率大的情况下,节点的度数必须很高,才能获得好的可靠性。另外,B. Fortz 等人针对 2-连通网络结构作了一系列研究,目标是以最小代价设计一个 2-连通网络,使得每条边所属的最短圈的长度都不超过约束长度,并且采用分支界定和启发式搜索方法求取该 NP 问题,取得了较好的效果。

IP 网络的拓扑结构需要和 IP 层运行的协议结合起来,才能保证 IP 网络的生存性。目前,IP 网络上常用的域内路由协议是链路状态路由协议,如 OSPF/IS-IS。运行这类协议的网络实际上是基于无连接的 IP 网络,由于其没有引进 MPLS 之类的面向连接的技术,因此称之为纯 IP 网络(pure IP network)。链路状态路由协议的自愈能力保证了纯 IP 网络的生存性:当网络局部出现失效时,链路状态路由协议在全网通告链路状态信息,各节点重算到其它节点的路由。只要整个网络连通,当路由协议收敛时,任意两个节点之间的通信仍然可以维持。但在协议收敛期间,可能形成环(micro-loop)。MPLS 网络由于其面向连接的特性,能够依靠较强的控制平面,将流切换到指定路径。这两种 IP 网络由于保护恢复的机制不同,其研究的侧重点也各不相同。

## 3 IP 网络的保护和恢复

### 3.1 纯 IP 网络保护恢复机制存在的问题

纯 IP 网络由于其逐跳转发的特性,使得在网络的保护恢复过程中存在以下 3 个问题:

(1)网络转发不连续。当网络局部出现失效时,OSPF 有两种方式发现拓扑变化:底层协议通过硬件检测到失效并报告给网络层协议;每隔 10 秒广播 Hello 报文,40 秒未收到邻居发来的 Hello 报文则说明相邻链路或者路由器失效。也就是说,发现失效的时间最多为 40 秒。另外,包括 LSA 报文的泛洪和 SPF 计算,重新路由后的收敛时间数量级在秒以上。在这段时间内,由于路由不能及时、正确地切换,造成大量的

丢包和服务不连续,对高可靠性要求的业务带来巨大的损失。

(2)转发连续性和路由稳定性之间存在矛盾。研究观测表明,网络中大多数的失效事件都是短时期的,时间数量级为几分钟。文<sup>[12]</sup>通过对 Sprint 骨干 IP 网络连续 5 个月的观测,对网络链路失效的时间规律进行了一定的研究。失效持续时间累积分布表明,大多数失效都是暂时性的;50%的失效情况持续时间小于 1 分钟,85%的失效情况持续时间小于 10 分钟。如果由硬件检测报告失效,由于收敛速度加快,短时间的失效造成连续的 LSA 泛洪和路由重算,从而在转发连续性和路由稳定性之间产生矛盾。

(3)网络性能下降。一方面,流量切换可能造成某些利用率比较大的链路拥塞,文<sup>[13]</sup>表明骨干网设计大多数时刻可以避免网络超负荷,80%的网络拥塞(链路利用率在 90%以上)是由短期链路失效(小于 1 分钟)造成的;另一方面,收敛过程中产生的环路可能造成网络性能急剧下降。下面以 Abilene 网络<sup>[14]</sup>为例说明环路的产生过程。图中链路上的数字代表链路权值。在网络正常状态下,从 Washington 到 Kansas, Denver, Sunnyvale, Seattle 的最短路径将经过 New York, Chicago, Indianapolis。假设 Indianapolis->Kansas City 的链路失效,那么 New York 将把到达 Kansas, Denver, Sunnyvale, Seattle 的数据包从 Washington 发送。如果 New York 比 Washington 早更新路由表,那么在两者更新的间隔期间,数据包将在两者之间形成环。

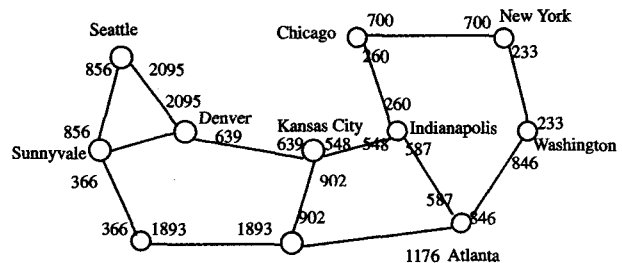


图 1 Abilene 网络的拓扑图

目前,不少学者围绕着以上几个问题展开了以下各个方面的研究,目的主要在于加快 IP 网络路由协议的收敛,快速切换流量,以尽量小的代价提高 IP 网络的生存性。

### 3.2 加快网络收敛速度

#### 3.2.1 网络失效检测

网络失效是网络连通性故障的一种重要表现,但网络失效也可能是在网络配置时人为造成的。网络生存性保障措施往往在检测到网络失效之后启动。加快网络失效检测速度可以减少路由器不可用时间、减少丢包率,提高网络的可用性和可靠性。

早在 2000 年,C. Alaettinoglu 等人就提出了 IGP 毫秒级收敛的草案<sup>[15]</sup>,提出有关加快网络失效检测速度的 3 条相关建议:毫秒级的 Hello 报文间隔,文<sup>[15]</sup>中的模拟实验表明,在 Hello 报文不是占用链路容量主要部分的情况下,减少 Hello 报文间隔对网络性能不会造成重要影响;LSP 报文转发在 SPF 计算之前完成;在数据包排队时赋予 Hello 报文高的优先级。

近年来,不少建议提出采用第二层检测和双向转发检测(Bi-directional Forwarding Detection, BFD)的机制来加快检测速度<sup>[16,17]</sup>。一般第二层网络都有自身的失效检测手段,但第二层的检测可靠性不强,设备驱动可能检测到了失效但不能通告路由器的操作系统。例如在一个交换域中,当交换机

下联的路由器失效时,交换机收到链路层失效指示信息,但没有机制可以将这一信息通告给网络中其它的路由器。BFD<sup>[18]</sup>是一种工作在两个转发引擎间的双向路径上的故障检测协议。IETF成立了专门的研究小组 BFD Working Group<sup>[19]</sup>,以加速 BFD 的发展进程。目前主流路由器厂家已经支持 BFD 协议。BFD 协议通过在点对点的链路之间定期(以 10 ms 为单位)发送故障检测包,不但可以检测和判断传输链路、光接口和设备端口等硬故障,还可以检测和判断传输层、链路层、IP 层和应用层由于误码造成的丢包等软故障,弥补了目前基于 SDH 故障检测只能实现传输层故障检测的不足。采用 BFD 后,故障检测时间在 30ms 左右。BFD 技术已经是新一代路由器端口故障检测的基本必备功能,不依赖于任何其它协议或者应用,由于采用硬件实现,不影响设备性能。采用 BFD 后,结合其他技术的进步,大型网络路由收敛时间可小于 500 ms,快速重路由时间可小于 50 ms。

无论是 Hello 报文还是第二层检测都不能区分链路失效和节点失效。当节点失效时,Hello 报文或者第二层检测都会判定相关的多条链路同时失效,不仅引发多个 LSA 报文泛洪,而且不利于各种快速重路由机制的使用。BFD 由于可以在两节点间的所有路径上发送探测报文,因此可以区分链路失效和节点失效,是今后作为网络失效检测的有力机制。

### 3.2.2 最短路经生成树的计算

典型的最短路经生成树(SPF)计算采用 Dijkstra 算法,计算时间为  $O(n \log(n))$ ,其中  $n$  为节点数目。增量式最短路经生成树(iSPF)算法<sup>[20]</sup>是一种有效的加速计算的方式,当网络拓扑发生变化时,iSPF 只对受影响的节点进行重新计算。文[15,16,21]都建议采用 iSPF 算法。

Sudip Misra 等人<sup>[22]</sup>则从随机网的角度出发,讨论如何计算一个统计意义上的最短路经生成树,使得网络局部失效时重新计算路由表的节点尽量少。Sudip Misra 等人采用的是自适应学习的方法,每当网络拓扑发生变化时都通过学习得到网络的统计状态。

### 3.2.3 路由信息表/转发表的更新和发布

文[16]的研究表明,路由信息表(RIB)和转发表(FIB)的更新是路由协议收敛的瓶颈,因为在一般的路由器结构中,RIB/FIB 的更新和 FIB 的发布共享 CPU。为了减小这个瓶颈的影响,一般是先更新和发布主要的前缀,然后以 50ms 为间隔交替更新和发布其它的前缀。一般,大型网络中需要 10 个多回合才能完成全部前缀的更新和发布。文[16]提出了三种加速 RIB/FIB 更新的机制:合并前缀;为前缀分配不同的优先权;优化路由表管理代码。

### 3.3 失效通告抑制

为了减小短期失效造成的不稳定,协议实现过程中往往在检测出失效后抑制 LSA 泛洪。Cisco 路由器配置中将这段抑制时间称为运载延迟(carrier delay),文[23]中建议该延迟为 2s。但失效通告抑制在长期的失效事件发生时会使网络收敛速度,因此必须调整 LSA 的抑制时间。文[15]建议分配给 link up 和 link down 事件以不同的抑制时间,使“坏消息”传播得快,“好消息”则被抑制一段时间再被传播。文[24]提出使用动态定时器调整 LSA 泛洪抑制时间。动态定时器的定时长短取决于网络的稳定状态:当网络稳定时,定时器定时短,网络拓扑发生变化时将在几毫秒内反应;当网络事件频繁发生时,定时时间呈指数增加。动态 LSA 泛洪抑制时间可以保证网络在稳定时得到快速的路由信息交换,在不稳定时

控制路由协议开销,提高网络稳定性。

### 3.4 本地快速重新路由

当检测到失效时,进行本地重新路由。将流量(deflection)切换到合适的下一跳,可以缩短网络转发不连续的时间。本地重新路由的关键是去除环路。故障不敏感路由 FIR<sup>[21,25~28]</sup>实际上就是一种本地重新路由的策略,以网络进口和目的节点作为参照自动切换流量,避免环路的产生。FIR 只需要对现有路由协议作比较小的改动,但在节点失效和多条链路同时失效的情况下性能不好,可能形成环路。IETF 的 Routing Area Working 小组(RTGWG)的工作也围绕着快速重路由协议的制定而展开,已经提交了一系列相关草案,涉及到的相关技术主要包括以下几点:

(1)无环替代<sup>[29]</sup>。如果链路  $I \rightarrow J$  失效,找到  $I$  的一个邻居  $N$ ,使得  $(I \rightarrow J) \notin \text{SPT}(N)$ , SPT 表示最短路径树,即流量全部切换到节点  $N$  后到达任意节点的流量都不会经过失效链路  $I \rightarrow J$ 。

(2)U-turns<sup>[30]</sup>。U 节点是  $I$  节点的一个邻居,如果  $U$  的一个邻居  $R$  满足  $(I \rightarrow J) \notin \text{SPT}(R)$ ,那么在链路  $I \rightarrow J$  失效时,流量通过  $I \rightarrow U \rightarrow R$  切换。

(3)IP 隧道<sup>[31]</sup>。在节点  $I$  计算两个集合  $A$  和  $B$ ,从  $I$  到集合  $A$  中的节点都不需要经过链路  $I \rightarrow J$ ,对于集合  $B$  中的任意节点  $b$ ,  $(I \rightarrow J) \notin \text{SPT}(b)$ ,  $A$  和  $B$  的交集即为可行的隧道末端节点。

(4)Notvia 地址<sup>[32]</sup>。给与节点  $I$  相连的  $J$  的接口另外绑定一个特定的地址  $JI$ ,在计算路由表时,计算到达  $JI$  的路径不能经过链路  $I \rightarrow J$ 。

以上技术都未进入实用阶段,各自都有相应的优缺点,文[33]对其进行了简要的评价。无环替代思路简单,但不能保证针对所有的失效链路  $I \rightarrow J$  都可以找到邻居  $N$ ,使得  $N$  到所有的目的地都不经过  $I \rightarrow J$ 。U-turns 机制通过选择两跳之内节点作为切换点,从而提高无环替代的链路保护范围,但节点必须具有区分正常数据流和被切换流的功能,这需要邻居节点间的协调和对网络接口的修改,实际操作比较困难。IP 隧道也不能使所有的链路得到保护,当找不到合适的隧道时,只能在  $A$  集合中选取一个潜在的元素  $a$  节点, $a$  可以采用无环替代找到相应的邻居  $N$ ,在切换过程中,将流量从隧道末节点转移到  $N$ 。但 IP 隧道仍然不能保证所有的链路都得到保护,并且需要某种 IP 封装机制的支持,如 IP-in-IP<sup>[34]</sup>, GRE<sup>[35]</sup>, L2TPv3<sup>[36]</sup>。Notvia 地址机制可以保护所有的链路和节点,但也需要某种 IP 封装机制,以支持 Notvia 地址,并且需要在全网内宣告 Notvia 地址。

不相交多路径路由(disjoint multipath routing)机制在流量均衡的同时也可以在一定程度上提高网络的生存能力——当某个节点或者链路失效时只会影响一部分流量。由于纯 IP 网络没有失效指示信号,因此不能将所有流量都切换到非故障路径上,但现有的不少快速重路由机制是借鉴多路径路由思想实现的。文[37]采用偏转(deflection)路由技术处理链路超负荷,同时可以用于处理链路失效事件,每个路由器做本地决策将流量从缺省出口导向特定出口。偏转路由的主要问题是不能产生环路。文[37]的研究表明,如果链路权重的比例在  $1+1/(d-1)$  以内,且网络中每个源、目的对之间有多条等长路径,那么可以提供无环偏转路径;并且进一步提供了对链路权值稍松弛的约束条件,保证得到的偏转路由是无环的。但是,该算法主要针对 PoP 结构的网络,当网络结构的格型

不够、不能保证 PoP 之间至少有两相等长度的路径时,不能使用该算法。文[38]提出 O2 算法,使某个节点到达目的节点的下一跳至少有两个。当检测到失效时选择备份的下一跳进行转发,但网络拓扑必须是格型结构并且具有足够的度数。

另外,不适用 MPLS 技术进行转发的 IP 网络实际上也可以利用 LSP 保护失效链路。当被保护链路出现失效时,直接通过 LSP 隧道重新路由。如果网络是双向的,LSP 隧道可以保护任意的失效链路,但需要网络对 RSVP-TE 的支持<sup>[39]</sup>。

### 3.5 网络链路权值调整

避免网络失效造成的超负荷一般很难只通过路由协议来完成。因为要实现流量的均衡需要对网络的流量有详细的了解,那么路由器必须保留相应的流信息,开销较大,现有的路由器结构并不支持。因此,一般从网络流量工程的角度出发,通过调整网络的配置疏导网络流量,实现流量均衡,基于链路权值调整的流量工程就是其中的方法之一。Bernard Fortz<sup>[40,41]</sup>等人以最小化加权链路利用率之和或最小化最大链路利用率为目标函数,采用 local search 的方法寻找最优的链路权值,并且在发生链路失效时采用 local search 的方法通过改变尽量少的链路权值,使得网络性能不会因为链路失效造成严重的下降。但在某些情况下,可能需要去掉某些链路,并且权值的改变是在最优化的结果基础上进行的。文[42~44]在权值设定之初就考虑可能出现的失效情况,使得网络在失效时也不会造成网络性能的急剧下降。这种方法只考虑出现的某条链路失效造成的影响或者关键失效影响的平均值,虽然避免了路由更新,但不可能覆盖全部可能的失效情况,而且计算量也比较大。文[45]提出了根据网络状态自动调整链路权值的框架,通过按序调整网络权值减少路由更新造成的网络不稳定的时间。

随着网络业务的发展,网络层路由协议的自愈能力已经不能满足纯 IP 网络的生存性需求。纯 IP 网络由于在网络层缺乏控制机制,快速高效的保护和恢复比 MPLS 要困难。但是将纯 IP 网络都升级为 MPLS 网络,显然是不现实的和不必要的。因此,纯 IP 网络的生存性研究目前受到了越来越多的关注。为了避免给现有网络带来大的影响,现有的研究思路主要是改进现有的路由协议,在其中添加或者修改尽量少的模块,以及调整各类网络参数,从而加快收敛和快速切换流量。

## 4 MPLS 网络的保护和恢复

纯 IP 网络的无连接特性给网络的保护恢复造成了许多困难,而面向连接的 MPLS 技术则为提高 IP 网络的生存性提供了有利的基础。MPLS 网络的故障检测可以通过链路连通性确认机制、MPLS ping 或者 RSVP Hello 协议来完成。MPLS 自愈技术则在避免 SDH50% 开销的情况下提供了高性能恢复的新选择,目前基于 MPLS 流量工程的路由切换对 IP 网络的生存性保证已基本可达到 1s。MPLS 自愈技术实现包括以下 3 种方式。

(1)链路或节点保护:这种方式与 SDH 的保护是类似的。为每个链路和节点提供单独的迂回路由进行保护,在建立标记交换路径时,每个节点负责为每条链路或节点计算保护路径。一旦某个链路或节点发生故障,立即被其直接上游节点检测到,然后在路由器上把流量立即切换到迂回路径。这种方法的优点是切换速度快,但缺点是需要很多备份资源,并

且路由器需要维护的状态也比较多。

(2)路由重计算:在标记交换路径发生故障后,通知该路径的入口路由器,该入口路由器再利用约束路由自动重新计算新的路由,并重新建立一条新的标记交换路径。由于有计算和重新建立过程,在实际网络中的恢复时间量级通常和路由恢复量级相当。它的主要优点是:无需手工安排额外的路径,直接和网络相适应。缺点在于恢复时间比较长。

(3)备份路径恢复:在建立标记交换路径时,指定其备份路径,在主路径发生故障时,通知入口路由器把流量切换到备份路径。它的主要优点是恢复时间比较快,主要缺点是需要占用额外的资源。

有关 MPLS 网络的保护恢复研究已开展了多年,最初的研究主要集中在 MPLS 保护和恢复路径的选取<sup>[46~49]</sup>。随着 MPLS 保护机制的日益成熟,近两年的研究主要围绕着恢复能力和服务质量的结合而展开。Achim Autenrieth<sup>[50~52]</sup>等人提出区分恢复 QoS(RD-QoS)框架,根据恢复时间将恢复的级别分为 4 类,对不同的恢复要求提供不同的 QoS 路由。Achim Autenrieth 还提到了扩展 RSVP、区分 QoS 服务,使其支持不同恢复等级。文[53]主要针对区分服务提出了 MPLS 恢复路由框架 CAIP,CAIP 扩散每条链路为各种类型的 LSP 和备份 LSP 预留的带宽以及剩余带宽。仿真实验表明,CAIP 比一般的 QoS 区分服务体系具有更好的保护恢复能力。文[54]介绍了一种将 QoS 恢复加入到现有 QoS 路由算法中的体系结构,其主要特点是通过工作路径中的链路数目、工作路径中可恢复的链路数目、需求带宽等多个参数计算 QoS 恢复度量(QoS Protection Metric),用于衡量保护恢复的需求,这种方式对于实际过程中量化保护恢复参数有很强的指导意义。另外,还有不少研究是针对满足保护需求和 QoS 的路径的选取而进行的。文[55]对离线流量工程生存性设计(OTESD)问题进行建模,目标是为每条流寻找受网络故障影响最小的主路径,同时满足带宽需求和优化资源利用率,并且提出了一种具有 4 个代价函数的启发式算法解决该问题。而文[56]则综合考虑主 QoS 路径和保护 QoS 路径集合,保护 QoS 路径集合由多段路径组成,分别保护主 QoS 路径上的某些部分,并且最小化整个网络的资源利用率。

影响 MPLS 自愈恢复保护性能的因素是路由器的负荷和需要入口路由器同时切换的标记交换路径数量。MPLS 自愈技术为 IP 网络的保护和恢复提供了快速灵活的策略,并且现有的 MPLS 保护恢复技术很容易移植到下一代的 GMPLS 网络中。但是基于 MPLS 的方法需要许多配置,从而可能引发人为的错误。更重要的是网络需要支持标记转发方式,但很多纯 IP 网络没有升级支持 MPLS,这也是目前纯 IP 网络的保护恢复受到越来越多关注的原因。

## 5 多层网络生存性的协调

由于网络的分层结构,每层都有自身的生存性机制,IP 层的生存性策略和其底层的生存性策略如果能够协调一致,将提高网络对故障的响应时间。对于在设计和管理流量工程机制中更好地优化网络性能也是很有价值的。多层网络生存性协调策略主要包括静态多层生存性策略和动态多层生存性策略。

静态多层生存性策略中,网络的拓扑结构不能改变,只能通过各层生存性维护的互联策略协调多层保护。互联策略通常由这样一些规则组成:它描述了什么时候开始和停止恢复

动作;如何协调不同生存性技术的动作顺序等。静态多层保护按照各层生存性策略的启动时序,分为从下到上和从上到下两种。从下到上的互联策略首先启动下层的恢复策略,然后通过 handoff 定时器<sup>[57]</sup>触发 IP 层的恢复策略。由于一般下层网络的恢复速度比 IP 层快,因此从下到上的恢复方式速度比较快,但效率不高。一方面,IP 层故障往往不能被低层检测,例如 MPLS 层报文出错而 SDH 检测不到,这时 IP 层必须在定时器规定时间之后才能启动恢复策略,而不能立即在检测到失效后就采取措施;另一方面,低层恢复策略涉及的粒度往往大于 IP 层,例如 MPLS 层可能只涉及到恢复一条流,但低层光网络就需要对整个光纤上复用的流进行切换。从上到下的互联策略则相反,由于 IP 层可以根据服务类型对各流采取恢复策略,所以从上到下的方式恢复粒度细,所用资源少,但 IP 层需要显示信号通知下层是否恢复,这在实际中比较难以应用,因此实际商用化的还是从下到上的互联策略。

多层网络的保护技术需要各层都预留一定的保护容量,网络预留容量的设计问题是静态多层保护必须解决的问题之一。现有的预留容量的设计方法主要有 3 种:重复保护(duplicated protection)、高层保护(logical spare unprotected)、共享池策略(common pool strategy)<sup>[58]</sup>。如果 IP 层逻辑链路 A 有保护逻辑链路 A',A'由低层的链路 B 承载,而低层的链路 B 又由 B'保护,那么 A 所承载的流量实际上得到了双重保护,这种预留容量的机制称为重复保护;如果没有为低层的链路 B 设计预留容量,那么就称为高层保护;如果低层的链路 B 和其它链路共享保护容量,则称为共享池策略。重复保护效率低、可靠性高,一般只在特殊的情况下使用,高层保护效率高但可靠性也相对低,共享池策略则是在两者之间的折衷。

对于骨干网来说,一般采用 IP-over-optical 的结构。为了增强控制功能,IETF 和 ITU 分别致力于 GMPLS 和 ASON 技术的研究和推广。GMPLS 和 ASON 虽然实现方式有所不同,但其目的都是加强多层网络的控制,建立智能化的光网络。在智能光网络中,控制平面可以启动安装和拆除光路,从而提供动态多层生存性策略。动态多层生存性策略通过改变光路,从而修改 IP 层逻辑拓扑,保证网络的正常恢复,避免在 IP 层预留保护带宽。动态多层生存性策略是一个新兴的研究对象,目前主要倾向于使用多层流量工程<sup>[58~61]</sup>的方式实现动态多层保护。多层流量工程结合了现有的流量工程技术和 GMPLS 自动光路切换技术,目标是提供服务质量保证。当网络状态发生变化时,多层流量工程不仅对 IP 流进行重新路由,而且可以通过重新配置逻辑拓扑以防止网络性能的恶化。多层流量工程正处于研究的初步阶段,重路由和逻辑拓扑重配置之间的联系还有待进一步探讨。

**总结** 网络的宽带化、移动性、网格化、泛在化催生了新的网络体系,下一代互联网(NGI)和下一代网(NGN)成为网络发展的趋势。NGI 和 NGN 在目标上有所不同,但都是基于 IP 技术构建的。因此,无论对于现在的互联网还是以后网络的发展,IP 网络的性能优化是保证互联网有效运行的核心。近年来,交互式应用和各种实时业务的增长对 IP 网络的生存性维护提出了更高的要求,有关 IP 网络的生存性也受到越来越多的关注。本文对 IP 网络生存性面临的问题以及影响因素作了概括,并且对目前有关纯 IP 网络生存性、MPLS 网络生存性、IP 网络与底层的生存性协调研究作了归纳总结,尤其对纯 IP 网络的生存性研究现状从多方面进行了比较详细的介绍。

保护恢复能力和服务质量在 MPLS 中的结合、动态多层保护是未来 IP 骨干网生存性维护的发展方向,已有越来越多的研究者和工作组投入到这两个方面的研究中来。

由于现有的网络大多还是基于纯 IP 技术构造的,因此纯 IP 网络的保护和恢复具有很好的实际意义。虽然纯 IP 网络的无连接特性给网络的保护恢复带来了一定的困难,但是在原有比较成熟的网络协议的基础上添加新的功能或者调整网络参数以增强 IP 网络的生存性是有实际意义的,现有的实验也证明是可行的。尤其是 IP 网络的快速重路由机制具有很好的潜力,许多方法和技术有待进一步研究。

## 参考文献

- Manchester J, Saha D, Tripathi S K. Guest editorial - Protection, restoration, and disaster recovery. *Network, IEEE*, 2004, 18(2):3~4
- Suwala G, Swallow G. SONET/SDH-like resilience for IP networks; a survey of traffic protection mechanisms. *Network, IEEE*, 2004, 18(2): 20~25
- Kirstaedter A. An Extended QoS Architecture Supporting Differentiated Resilience Requirements of IP Services, draft-kirstaedter-extqosarch-00, txt, July 2000
- Pei Dan, Massey D, Zhang Lixia. A Framework for Resilient Internet Routing Protocols. *Network, IEEE*, 2004, 18: 5~12
- Iannaccone G, Chuah C, Bhattacharyya S, et al. Feasibility of IP restoration in a tier-1 backbone. *IEEE Network Magazine*, January-February 2004
- Dubios N, Fondeviole B, Michel N. Fast convergence project, January 2004. <http://www.ripe.net/ripe/meetings/ripe-47/presentations/ripe47-routing-fcp.pdf>
- Awduche D, Chiu A, Elwalid A, et al. Overview and principles of internet traffic engineering[S/OL]. 2002-05, IETF RFC3272
- Autenrieth A, Kirstaedter A. Fault Tolerance and Resilience Issues in IP-based Networks. In: Second International Workshop on the Design of Reliable Communication Networks, München, April 2000
- Sharma V. Framework for Multi-Protocol Label Switching MPLS-based Recovery. RFC3469, February 2003
- Kim Minkyu, M'edard M. Robustness in Large-Scale Random Networks. In: INFOCOM 2004, Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 4, March 2004. 2364~2373
- Weichenberg G E, Chan V W S, M'edard M. High-Reliability Architectures for Networks Under Stress. In: INFOCOM 2004, Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 1, March 2004. 131~141
- Iannaccone G, Chuah Chen-Nee, Mortier R, et al. Analysis of link failures in an IP backbone. In: Proc. ACM Sigcomm Internet Measurement Workshop, Nov. 2002
- Iyer S, Bhattacharyya S, Taft N, et al. An approach to alleviate link overload as observed on an IP backbone. In: INFOCOM 2003, Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE, vol 1. 30 March-3 April 2003. 406~416
- <http://abilene.internet2.edu/>
- Alaettinoglu C, Jacobson V, Yu H. Towards Milli-Second IGP Convergence, November 2000. draft-alaettinoglu-ISIS-convergence-00
- Francois P, Filis C, Evans J, et al. Achieving Sub-Second IGP Convergence in Large IP Networks. *ACM SIGCOMM Computer Communication Review*, 2005, 35(2)
- Pasqualini S, Iselt A, Kirstaedter A. MPLS Protection Switching vs. OSPF Rerouting - A Simulative Comparison. In: Fifth International Workshop on Quality of future Internet Services (QofIS'04), Barcelona, Spain, September 2004
- Katz D, Ward D. Bidirectional Forwarding Detection, draft-ietf-

- bfd-base-04. txt, October 2005
- 19 Bidirectional Forwarding Detection (bfd) Charter. <http://www.ietf.org/html.charters/bfd-charter.html>
  - 20 McQuillan J M, Richer I, Rosen E C. An overview of the new routing algorithm for the arpanet. In: Proceedings of the Sixth Symposium on Data Communications, ACM Press, 1979. 63~68
  - 21 Nelakuditi S, Lee Sanghwan, Yu Yinzhe, et al. Failure Insensitive Routing for Ensuring Service Availability. In: Proc. International Workshop on Quality of Service (IWQoS), 2003
  - 22 Misra S, Oommen B J. GPSPA a new adaptive algorithm for maintaining shortest path routing trees in stochastic networks. International Journal of Communication Systems, 2004. 963~984
  - 23 Iannaccone G, Chuah Chen-Nee, Mortier R, et al. Analysis of link failures in an IP backbone. In: Proc. ACM Sigcomm Internet Measurement Workshop, Nov. 2002
  - 24 Martey A. IS-IS Network Design Solutions. Cisco Press, 2002
  - 25 Lee Sanghwan, Yu Yinzhe, Nelakuditi S, et al. Proactive vs. Reactive Approaches to Failure Resilient Routing. In: INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies, Vol 1, March 2004. 176~186
  - 26 Zhong Zifei, Nelakuditi S, Yu Yinzhe, et al. Failure Inferencing-based Fast Rerouting for Handling Transient Link and Node Failures. In: INFOCOM 2005, 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Vol 4, March 2005. 2859~2863
  - 27 Nelakuditi S, Lee Sanghwan, Yu Yinzhe, et al. Fast Local Rerouting for Handling Transient Link Failures. IEEE/ACM Transactions on Networking. <http://arena.cse.sc.edu/papers/fir.ton04.pdf>
  - 28 Zhong Zifei, Keralapura R, Nelakuditi S, et al. Avoiding Transient Loops through Interface-Specific Forwarding. In: 13th International Workshop on Quality of Service - IWQoS 2005, Passau, Germany, June 2002. 19~232
  - 29 Atlas A, Zinin A. Basic Specification for IP Fast-Reroute; Loop-free Alternates. draft-ietf-rtgwg-ipfrr-spec-base-05, February
  - 30 Atlas A. U-turn alternates for IP/LDP Local Protection. draft-atlas-ip-local-protect-uturn-00. txt, November 2004
  - 31 Bryant S, Filsfils C, Previdi S, et al. IP Fast Reroute using tunnels. draft-bryant-ipfrr-tunnels-01, Oct. 2004
  - 32 Bryant S, Shand M. IP Fast Reroute Using Notvia Addresses. draft-bryant-shand-ipfrr-notvia-addresses-00, March 2005
  - 33 Francois P, Bonaventure O. An Evaluation of IP-based Fast Reroute Techniques. In: CoNEXT'05, October, Toulouse, France, 2005
  - 34 Simpson W. RFC-1853, IP in IP Tunneling. RFC1853, October 1995
  - 35 Hanks S, et al. Generic Routing Encapsulation (GRE). RFC1701, October 1994
  - 36 J, et al. Layer Two Tunneling Protocol (Version 3). draft-ietf-l2tpext-l2tp-base-14. txt, June 2004
  - 37 Iyer S, Bhattacharyya S, Taft N, et al. An approach to alleviate link overload as observed on an IP backbone. In: INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE, vol 1, 30 March-3 April 2003. 406~416
  - 38 Schollmeier G, Charzinski J, Adter A K, et al. Improving the resilience in IP networks. In: Workshop on High Performance Switching and Routing, HPSR, June 2003. 91~96
  - 39 Pan P, Swallow G, Atlas A. Fast Reroute Extensions to RSVP-TE for LSP Tunnels. RFC4090, May 2005
  - 40 Fortz B, Thorup M. Internet Traffic Engineering by Optimizing OSPF Weights. In: INFOCOM 2000, Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Vol 2, March 2000. 519~528
  - 41 Fortz B, Thorup M. Optimizing OSPF/IS-IS Weights in a Changing World. IEEE Journal on Selected Areas in Communications, 2002, 20(4)
  - 42 Fortz B, Thorup M. Robust optimization of OSPF IS-IS Weights. In: Proc. INOC 2003, October 2003. 225~230
  - 43 Nucci A, Schroeder B, Bhattacharyya S, et al. IGP Link Weight Assignment for Transient Link Failures. [ Technical Report ]. TR02-ATL-071000. SPRINT ATL. <http://cambridgeweb.cambridge.intel-research.net/people/pub/cdiot/TR02-ATL-071000.pdf>
  - 44 McQuillan J M, Richer I, Rosen E C. An overview of the new routing algorithm for the Arpanet. In: Proceedings of the Sixth Symposium on Data Communications, ACM Press, 1979. 63~68
  - 45 Gous A, Afrakhteh A. Traffic Engineering through automated optimization of routing metrics. In: Trans-European Research and Education Networking Association, Rhodes, Greece, June 2004
  - 46 Ahn G, Jang Jongsoo. An Efficient Rerouting Scheme for MPLS-Based Recovery and Its Performance Evaluation. Telecommunication Systems, 2002. 481~495
  - 47 Bremler-Barr A, Afek Y, Kaplan H, et al. Restoration by path concatenation; Fast recovery of MPLS paths. In: Proc. ACM Symposium on Principles of Distributed Computing, 2001. 43~52
  - 48 Otel F. On fast computing bypass tunnel routes in MPLS-based local restoration. In: Proceedings of 5th IEEE International Conference on High Speed Networks and Multimedia Communications, Jeju, Korea, 2002. 234~238
  - 49 Bartos R, Raman M. A scheme for fast restoration in MPLS networks. In: Proc. of the Twelfth IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS), 2000. 488~493
  - 50 Autenrieth A, Kirstädter A. Resilience-Differentiated QoS - Extensions to RSVP and DiffServ to Signal End-to-End IP Resilience Requirements. In: Third International Workshop on the Design of Reliable Communication Networks, Budapest, Hungary, October 2001
  - 51 Autenrieth A. Recovery Time Analysis of Differentiated Resilience in MPLS. Design of Reliable Communication Networks (DRCN) 2003, Banff, Alberta, Canada, October 333~340
  - 52 Autenrieth A, Kirstädter A. RD-QoS; The Integrated Provisioning of Resilience and QoS in MPLS-based Networks. In: IEEE International Conference on Communications, New York, USA, May 2002
  - 53 Dogar F R, Uzmi Z A, Baqai S M. CAIP: A Restoration Routing Architecture for DiffServ Aware MPLS Traffic Engineering. In: Proc. of 5th International Workshop on Design of Reliable Communication Networks (DRCN 2005), Oct. 2005. 55~60
  - 54 Marzo J L, Calle E, Scoglio C, et al. Adding QoS Protection in Order to Enhance MPLS QoS Routing. In: IEEE International Conference on Communications, Vol 3, May 2003, 2003. 1973~1977
  - 55 Amin M, Ho Kin-Hon, Pavlou G, et al. Improving Survivability Through Traffic Engineering in MPLS Networks. In: Proc. of Computers and Communications, June 2005. 758~763
  - 56 Bejerano Y, Breitbart Y, Orda A, et al. Algorithms for Computing QoS Paths with Restoration. IEEE/ACM Transactions on Networking, 2005, 13(3): 648~661
  - 57 Vasseur J P, et al. Network Recovery, Protection and Restoration of Optical, SONET-SDH, IP and MPLS, Norgan Kanfmann Series in Networking, Elsevier, 2004
  - 58 Pickavet M, Demeester P, Colle D, et al. Recovery in Multilayer Optical Networks. Journal of Lightwave Technology, January 2006, 24(1): 122~134
  - 59 Puype B, Vasseur J-P, Groebbens A, et al. Benefits of GMPLS for Multilayer Recovery. IEEE Communications Magazine, July 2005. 51~59
  - 60 Yan Qiang, Colle D, De Maesschalck S, et al. Performance Evaluation of Multi-Layer Traffic Engineering Enabled IP-over-ION Networks. Photonic Network Communications, 2005. 255~280
  - 61 Vigoureux M, Berde B, et al. Multilayer Traffic Engineering for GMPLS-Enabled Networks. IEEE Communications Magazine, July 2005. 44~50