

可重构服务承载网愈合机理研究

缪宇霆¹, 吴春明¹, 杨强², 姜明³

(1. 浙江大学 计算机科学与技术学院, 浙江 杭州 310027; 2. 浙江大学 电气工程学院, 浙江 杭州 310027;
3. 杭州电子科技大学 计算机科学与技术学院, 浙江 杭州 310027)

摘要: 提出了一种服务承载网的快速愈合算法。该算法将服务承载网的愈合问题等价地转化为多商品流问题进行解决。通过大量的仿真实验的验证和量化分析表明, 该方法相比较传统的全局重映射方法, 能更高效地提高服务承载网的愈合成功率以及愈合效率, 从而快速地实现网络服务的恢复, 并可显著地降低对底层网络负载的影响。

关键词: 网络虚拟化; 服务承载网; 抗毁; 重构; 愈合

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2012)08-0052-10

Self-healing mechanism for reconfigurable service overlay networks

MIAO Yu-ting¹, WU Chun-ming¹, YANG Qiang², JIANG Ming³

(1. College of Computer Science, Zhejiang University, Hangzhou 310027, China;
2. College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China;
3. College of Computer Science, Hangzhou Dianzi University, Hangzhou 310027, China)

Abstract: A fast self-healing algorithm of reconfigurable service overlay networks against the physical substrate faults was presented. The suggested algorithm converted the self-healing problem to the multi-commodity flow problem. Extensive simulation experiments and the numerical results show that the proposed fast self-healing algorithm can further improve the healing success ratio and efficiency of the service overlay networks in comparison with the conventional overall re-mapping approach, whilst significantly reduces the load across the physical network.

Key words: network virtualization; service overlay network; survivability; reconfiguration; healing

1 引言

当今的互联网信息产业正处于高速发展状态。在传统的网络体系结构下, 依靠拓展链路传输带宽, 提高节点处理能力, 增加节点存储容量, 融入网络控制协议等升级扩展技术已经难以满足日益扩大的用户业务承载需求, 并且随之产生了一系列网络复杂度快速提高和数据传输效率显著降低的

问题。可重构柔性网络^[1-3]——一种面向服务提供的新型网络体系结构的出现, 可以使互联网摆脱传统网络技术的束缚, 解决当前互联网面临的 2 大问题: 1) 网络是刚性的, 只能依靠升级和扩展; 2) 网络资源是共享的, 彼此之间会互相干扰。

在可重构柔性网络中, 用户业务与网络服务的关系从传统的紧耦合变为松耦合: 所构建网络提供服务的方式是基于网络自身的能力, 而不是特定的

收稿日期: 2011-08-30; 修回日期: 2012-04-13

基金项目: 国家重点基础研究发展计划(“973”计划)基金资助项目(2012CB315903); 国家自然科学基金资助项目(61070157, 61070213, 61103200); 浙江省重点科技创新团队基金资助项目(2011R50010)

Foundation Items: The National Basic Research Program of China (973Program) (2012CB315903); The National Natural Science Foundation of China (61070157, 61070213, 61103200); Key Science and Technology Innovation Team Project of Zhejiang Province (2011R50010)

用户业务需求,从而实现了用户与服务分离;每种网络服务可以支持多种特性相似的用户业务,新的业务可以利用原有的网络服务作为支撑,使得网络需要升级的几率大幅降低。

可重构柔性网络技术利用构建可重构服务承载网的方式来提供相应的网络服务。在服务承载网中,网络服务提供的质量直接影响用户需求的满足程度和用户请求的响应速率。因此,服务承载网需要在底层网络发生故障时快速愈合,才能尽可能快地恢复正常的网络服务供应,维持终端用户继续正常地获取他们所订阅的服务。

本文研究了服务承载网的快速愈合机理并提出了服务承载网的快速愈合算法。该算法区别于传统的对失效服务承载网进行全局重映射的方法,它通过引入多商品流问题^[4](multi-commodity flow problem)的解决策略,将服务承载网愈合的问题转化为多商品流问题,通过解一次多商品流问题愈合多个损坏的服务承载网,以实现承载网发生故障时的快速愈合,从而使得服务承载网具有快速可重构这一特性。该方法可以提高服务承载网的愈合成功率,降低愈合时的时间消耗,并且可以保持整个底层网络在服务承载网愈合之后的负载均衡。

2 相关工作

通常,服务承载网可被认为是一种面向不同用户提供差异化服务的虚拟网络,用于满足用户对服务的个性化需求。利用网络虚拟化技术可构建多个提供异质服务的承载网,并同时共享底层物理网络的资源。因此,可以认为网络虚拟化技术是实现服务承载网构建、运行和维护的重要技术实现方式。

众所周知,目前基于网络虚拟化技术的服务承载网构建问题大部分是 NP 完全(NP-complete)问题。近年来启发式算法和线性规划方法的出现^[5~10],初步解决了此类问题存在计算复杂度高的问题。目前,网络虚拟化技术在服务承载网构建方面的应用研究仍集中在如何实现网络映射和优化资源利用方面,且已存在较多成熟机制与算法,但对于在故障情形下的服务承载网愈合问题尚缺乏有效手段。已有的用于解决服务承载网愈合问题的解决方案仍采用全局重新映射的方法。该方法利用文献[7]中提出的构网算法,将失效的服务承载网依次进行节点重映射和链路重映射。其愈合效果等效于将整

个服务承载网在底层物理网络之上进行一次重新映射。这将导致服务承载网愈合速度慢,且将严重影响和干扰其他虚拟网的正常运行和端到端的服务提供。

本文认为,可重构服务承载网的快速愈合问题既是承载网重构(重映射)问题,同时也与其自身抗毁性相关。近几年来,虽然服务承载网抗毁方面的研究也取得了较多成果^[11~13],但仍然存在诸多不足之处。文献[11]提出了一种基于分布式 Agent 的虚拟资源自管理机制,用于监控运行于底层网络上方的服务承载网是否发生故障,并且也提出了承载网“愈合”的概念,但是并没有给出具体的愈合方法。文献[12]设计出了一种基于共享备用资源的服务承载网保护方法,该方法按照预先给每个承载网分配一定数量的备用网络资源,用于避免在底层网络发生故障时带来的承载网失效问题。这种方法可以看作是一种离线的服务承载网保护方法,不能很好地解决在线的承载网失效问题。文献[13]给出的网络环境中错误诊断的策略主要基于及时发现并解决底层网络中出现的硬件故障的方法来恢复失效的服务承载网。当某些硬件故障受客观条件限制不能被及时修复时,就会降低服务承载网恢复的效率。

总的来说,成熟的服务承载网愈合方法还未被研究,现存的愈合方法基本上等价于传统的构网算法,即在底层网络发生故障时对失效的服务承载网进行全局重映射;此外,服务承载网抗毁性方面的研究大部分又基本上是基于共享资源保护方面的策略,并未具体实现服务承载网的愈合方法。本文针对服务承载网愈合方法研究的不足,提出了基于多商品流问题思想的可重构服务承载网快速愈合算法,旨在提高服务承载网的愈合能力,加快服务承载网在遭遇故障时的愈合速率,以及缩小服务承载网在愈合时的重构规模。

3 问题描述

随着网络虚拟化研究的深入,网络映射问题的定义已经变得非常成熟。本节给出了物理网络,服务承载网以及服务承载网映射问题的公式化定义,以便在第4节更加清晰地描述可重构服务承载网的快速愈合机制。为了方便起见,本文采用文献[8]中提出的网络映射的问题描述。

3.1 物理网络

本文用无向图来给所用到的物理网络进行建

模, 定义为 $G^p = (N^p, L^p, C_N^p, C_L^p)$, 其中, N^p 和 L^p 分别表示物理节点和物理链路的集合。对于任意物理节点 $n \in N^p$, 节点的能力 C_N^p 表示物理节点的主要属性; 同时, 对于任意物理链路 $l \in L^p$, 链路的带宽 C_L^p 表示物理链路的主要属性。

3.2 服务承载网

本文同样用无向图来给所用到的服务承载网进行建模, 定义为 $G^v = (N^v, L^v, C_N^v, C_L^v)$, 其中, N^v 和 L^v 分别表示虚拟节点和虚拟链路的集合。对于任意虚拟节点 $n \in N^v$, 节点所需求的能力 C_N^v 表示虚拟节点的主要属性; 同时, 对于任意虚拟链路 $l \in L^v$, 链路所需求的带宽 C_L^v 表示虚拟链路的主要属性。

3.3 服务承载网到物理网络的映射

服务承载网 G^v 到物理网络 G^p 的映射 M 可以定义如下: 找到一个 G^p 的子图 $G^{p'} = (N^{p'}, L^{p'})$, 使得所有的虚拟节点都被映射到不同的物理节点, 并且所有的虚拟链路都被映射到由若干物理链路构成的无环物理路径, 即:

$$M : M(G^v) = (N^{p'}, P^{p'}) \quad (1)$$

其中, $N^{p'}$ 表示物理节点的子集, $P^{p'}$ 表示物理网络中的无环物理路径的集合。 M 可以进一步分解为节点映射以及链路映射:

$$M_N : M_N(N^v) = N^{p'} \quad (2)$$

$$M_L : M_L(L^v) = P^{p'} \quad (3)$$

若 $M(G^v)$ 满足以下 2 个限制条件:

$$C_N^v \leq C_N^p, \forall n \in N^v, M_N(n) \in N^{p'} \quad (4)$$

$$C_L^v \leq C_L^p, \forall l \in L^v, M_L(l) \in P^{p'} \quad (5)$$

则 M 是一个合法的网络映射。

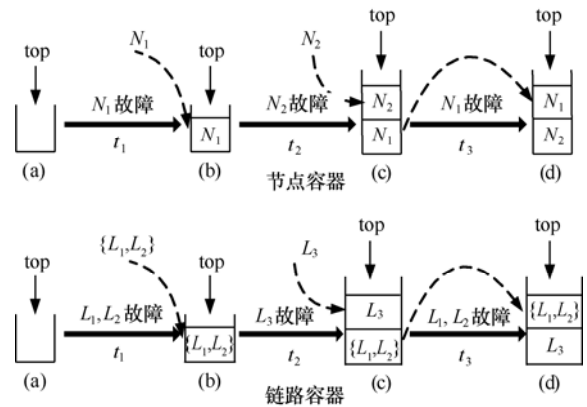
4 可重构服务承载网的快速愈合机制

服务承载网在满足用户需求的程度和响应用户请求的速率这 2 个方面都有着极高的要求。因此, 服务承载网必须在底层网络发生故障时快速愈合, 才能尽可能快地恢复正常的网络服务供应, 从而使终端用户正常地获取他们所订阅的服务。在研究可重构服务承载网快速愈合机制的过程中, 仅仅设计出一种服务承载网的快速愈合算法还不能满足复杂多变的网络环境, 必须加入一套辅助快速愈合算法运行的机制, 才能保证可重构服务承载网快速愈合算法达到迅速、可靠、高效的预期目标。本节将

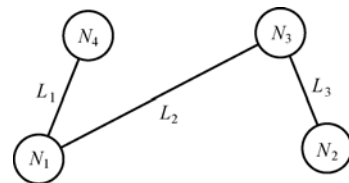
详细介绍可重构服务承载网快速愈合机制的 3 个重要组成部分: 底层物理网络设备可靠性的评估方法, 服务承载网的快速愈合算法和未能实时愈合的服务承载网的排队策略。

4.1 物理网络设备可靠性评估

服务承载网快速愈合的问题在某种程度上也可以看作是网络重映射的问题, 目的是把一些映射在发生故障的物理设备(物理节点、链路)之上的虚拟节点、链路重映射到正常工作的物理设备。如果在进行重映射之后, 无法保障这些新的用于支持虚拟节点、链路的物理设备的可靠性, 那势必会影响到服务承载网愈合后的正常工作。因此, 必须尽可能地把这些失效的虚拟节点、链路映射到可靠性高的物理设备之上, 以保证服务承载网愈合的稳定性。为此, 本文设计了 2 个集中式的容器用来分别记录整个底层网络中曾经发生过故障的物理节点和物理链路, 称之为节点容器和链路容器, 并采用基于“局部性原理(locality principle)”的“最近最少故障(least recent failed)”策略来评估这些物理设备的可靠性。如图 1 所示, 该策略规定, 最近一次发生故障的物理设备的标识符会被记录在容器顶端(图中标记为 top)。



(a) 物理设备可靠性评估容器



(b) 底层网络拓扑

图 1 可靠性评估示例

图 1(a)给出了图 1(b)中所示的底层网络的节点容器和链路容器, 初始状态均为空。在 t_1 时刻, 节

点 N_1 发生故障，根据评估策略， N_1 的标识符被加入到节点容器的顶端，与 N_1 相连的链路 L_1, L_2 受到 N_1 故障的影响而无法正常使用，所以被对应地加入到链路容器的顶端；在 t_2 时刻，节点 N_2 第一次发生故障，因此 N_2 的标识符被加入节点容器中并被记录在顶端，与 N_2 相连的链路 L_3 也被加入到链路容器的顶端， L_1, L_2 相应地下降一个位置； t_3 时刻的情况类似 t_1 ，节点容器顶端被 N_1 的标识符占据，因为 N_1 又发生了一次新的故障，同时 L_1, L_2 也被移动到链路容器的顶端。

物理节点 n 或物理链路 l 的可靠性可以按如下方法定义：

$$r_n = \frac{pos[n]}{|V|+1}, r_l = \frac{pos[l]}{|E|+1} \quad (6)$$

其中， $|V|$ 和 $|E|$ 分别表示物理网络中节点和链路的总数， $pos[n], pos[l]$ 代表物理节点 n 或物理链路 l 在各自容器中的位置索引。如图 1(a) 所示，物理节点 N_1 的位置索引为 1， N_2 的位置索引为 2，即 $pos[N_1]=1, pos[N_2]=2$ ；物理链路 L_1, L_2 的位置索引均为 1， L_3 的位置索引为 2，即 $pos[L_1]=pos[L_2]=1, pos[L_3]=2$ 。根据式(6)可知， $r_{N_1}=0.2, r_{N_2}=0.4, r_{L_1}=r_{L_2}=0.25, r_{L_3}=0.5$ 。显然，如果某个物理节点或某条物理链路从未发生过故障，那么它也不可能被记录在该容器中，此时本文规定这样的物理设备的可靠性固定为 1。

4.2 服务承载网的快速愈合算法

通常来说，底层网络节点或链路的故障会导致当前的网络映射（虚拟节点、链路映射）不能正常使用，从而发生服务承载网无法正常提供服务的情况。本文的核心工作是研究失效服务承载网的快速愈合算法。该算法会尽可能多地保留原有的可用网络映射，从而避免将失效的服务承载网进行全局重映射。首先讨论由物理链路故障造成服务承载网失效的问题，再讨论由物理节点故障造成服务承载网失效的问题。因为物理节点的故障同时也会导致所有连接到它的物理链路也发生故障。

4.2.1 物理链路故障

当承载虚拟链路的物理链路发生故障时，这些虚拟链路所对应的服务承载网就会无法正常工作。应对这种情况最直接的方法是找出所有因链路故障而发生状况的服务承载网，并将这些承载网依次进行重映射，包括重映射每个承载网的所有虚拟节

点和虚拟链路，并完全忽略那些没有受影响的虚拟节点、链路映射。值得注意的是，在通常情况下，大多数发生状况的服务承载网中的虚拟节点、链路映射依然可以正常工作。如果底层网络的管理者把这些可用的网络映射先保存起来用于重用，将是一个不错的选择。更重要的是，虚拟链路通常是被映射在一些由连续的物理链路构成物理路径之上。这意味着虚拟链路通常只是被部分映射在某条物理链路之上。因此，单个物理链路故障极有可能只是引发了用于承载虚拟链路的物理路径的某一段不能正常使用。如图 2 所示，物理链路 CD 发生故障导致虚拟链路 \overline{ab} 映射在其上方的片段不能正常工作，而映射在物理链路 \overline{AB} 和 \overline{BC} 上的部分则不受影响。事实上，一个失效的服务承载网中的大部分网络映射是可重用的，这也为提高服务承载网的愈合的效率，实现其可重构的特点提供了可能。

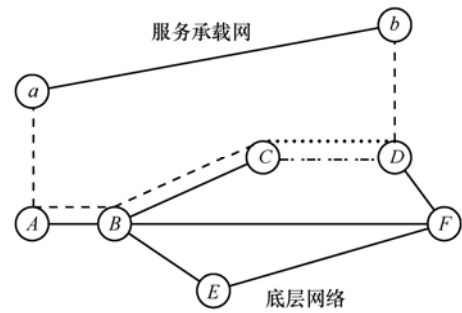


图 2 虚拟链路部分故障

假设有一条物理链路 $L^p = \overline{AB}$ 发生故障，有 k 条虚拟链路映射在 \overline{AB} 上，记为 $L_1^v, L_2^v, \dots, L_k^v$ 。这 k 条虚拟链路分别包含在 r 个不同的服务承载网中，记为 $V_1, V_2, \dots, V_r (1 \leq r \leq k)$ ($r \leq k$ 是因为同一服务承载网的虚拟链路可以映射到相同的物理链路)。物理链路 \overline{AB} 的故障会导致这 k 条虚拟链路的映射变得不可用，同时这 r 个服务承载网也将不再正常工作。从网络流的角度来看，这意味着物理节点 A 和 B 之间损失了 k 个流量大小分别为 $f_1^v = C_1^v, f_2^v = C_2^v, \dots, f_k^v = C_k^v$ 的商品流 ($C_1^v, C_2^v, \dots, C_k^v$ 的定义见第 3 节)。这样 r 个服务承载网的愈合问题就等价地转化为解决物理节点 A 和 B 之间总流量为 F ，并满足以下限制条件的多商品流问题：

$$\begin{cases} F = \sum_{i=1}^k f_i^v \\ f_1^v = C_1^v, f_2^v = C_2^v, \dots, f_k^v = C_k^v \end{cases} \quad (7)$$

在得到一组可行解之后，只需要把不同的商品流所经过的物理路径加入到它们所对应的服务承载网映射中，即可完成服务承载网的愈合。

利用多商品流问题解决服务承载网的愈合问题可以有效地在愈合过程中减轻底层网络的负载，并且显著地提高服务承载网愈合的成功率，因为这种方法允许将一条虚拟链路映射到多条物理路径之上^[8]。更重要的是，以上 r 个服务承载网的愈合只需要通过解一次多商品流问题即可解决，这也在很大程度上提高了服务承载网的愈合效率。

服务承载网愈合问题等价地转化为多商品流问题解决还需要考虑以下 3 个方面的问题。

1) 数据分组的乱序: 多商品流问题的解决方案允许将一条虚拟链路映射到多条物理路径之上，从而带来了数据分组从源到汇传输的过程中发生乱序(out-of-order)的可能。所需要做的额外工作是在愈合机制中加入一些防止数据分组在进行传输时发生乱序的控制方案^[14-16]。

2) 可行解不存在: 如果当前网络中的资源紧张，不存在满足式(7)中限制的多商品流可行解时，可以选择减小总流量 F 的方法，从而先愈合部分服务承载网。底层网络管理者可以根据服务承载网所提供网络服务的类型，适当减小总流量 F 的大小。例如，若服务承载网 V_1 和 V_2 提供的是非实时响应的服务，则可以从 F 中去掉 f_1^v 和 f_2^v 这 2 个流量。在 F 减小之后，再用多商品流问题的解愈合剩下的服务承载网。

3) 流的选择: 利用多商品流解决服务承载网的愈合问题有可能产生多组可行解，即在源和汇之间存在多种传输商品流的方案。这时需要选择可靠性最高的一种。假设一个商品流的传输方案需要经过 n 个物理节点和 l 条物理链路，那么该商品流传输方案的可靠性 φ 定义为

$$\varphi = \prod_{i=1}^n r_i \prod_{j=1}^l r_j \quad (8)$$

其中， r_i 和 r_j 分别代表物理节点和链路的可靠性(式 6)，其中， r_i 和 r_j 的大小均不大于 1。式(8)意味着如果多商品流问题的解所经过的物理节点和链路越少，那么这条商品流的可靠性就越高。这是因为如果商品流传输路线上的任意一个节点或链路发生故障，这条商品流就不可能达到预期的流量，所以它所经过路线上的任何一个节点或链路都不能

发生故障，才能保证这条商品流的正常传输。

服务承载网 v_1, v_2 各虚拟链路的需求带宽以及底层网络 PN 的物理链路带宽如图 3 所示。为更方便地描述服务承载网的愈合过程，图中不考虑物理节点的能力值 C_N^p 。物理链路 \overline{BF} 发生故障引发 v_1, v_2 的虚拟链路映射 $M_1(\overline{ab}) = \overline{ABF}$, $M_2(\overline{cd}) = \overline{BFG}$ 不可用。服务承载网快速愈合算法会在节点 B 和 F 之间解一个流量大小为 $F = f_1 + f_2 = C_{\overline{ab}} + C_{\overline{cd}} = 7$ 的多商品流问题，并得到一组可行解：流量为 3 的流 f_1 通过物理路径 \overline{BCDF} 传输；流量为 4 的流 f_2 通过物理路径 \overline{BCDF} 和 \overline{BEF} 传输，这 2 条物理路径各自传输的流量大小均为 2。最后更新与这 2 个商品流对应的虚拟链路映射，得到 $M_1(\overline{ab}) = \overline{ABCDF}$, $M_2(\overline{cd}) = \{\overline{BCDFG}, \overline{BEFG}\}$ 。至此，服务承载网 v_1, v_2 愈合完成。

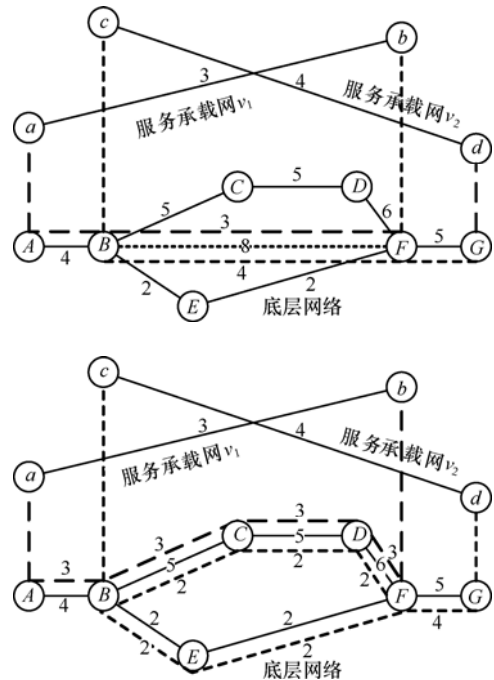


图 3 链路故障引发的服务承载网愈合

4.2.2 物理节点故障

在讨论完由物理链路故障导致服务承载网失效及其愈合算法之后，服务承载网因物理节点故障发生失效，并进行愈合的问题就变得相对简单。概括地讲，只需要把那些映射在故障物理节点上的虚拟节点转移到其他合适的物理节点，再更新那些因重映射虚拟节点而发生变化的虚拟链路映射即可完成服务承载网的愈合。

选择合适的物理节点进行虚拟节点重映射需要考虑以下 3 个方面的问题。

1) 可靠性: 新物理节点的可靠性直接影响到服务承载网愈合的过程可靠程度。如果新的物理节点经常发生故障, 则需要再次更换物理节点用作虚拟节点的重映射。

2) 资源: 应该尽可能地把虚拟节点转移到那些拥有更多资源的物理节点上, 这样可以有效的保持底层网络的负载均衡, 并且提高网络资源的利用率。物理节点 n 的资源 R_n 定义如下:

$$R_n = C_n^p \sum_{l \rightarrow n} C_l^p \tag{9}$$

其中, $l \rightarrow n$ 表示物理链路 l 与物理节点 n 相邻接, C_n^p 与 C_l^p 的定义见 3.3 节。

3) 距离: 新的物理节点与发生故障的物理节点的距离对服务承载网的愈合也存在一定的影响。如果距离发生故障的节点太近, 有可能会存在可用物理链路过少的隐患; 如果距离发生故障的节点太远, 则在服务承载网愈合之后, 会存在数据分组传输距离过远的问题。

综合 1) 和 2), 定义物理节点的价值:

$$\psi = R_n \sum_{l \rightarrow n}^{(m \sum r_l)} \tag{10}$$

其中, $l \rightarrow n$ 表示物理链路 l 与物理节点 n 相邻接, r_n 与 r_l 的定义见 4.1 节; 根据 3), 本文限定新物理节点的选择范围 H , H 代表新物理节点与距离故障物理节点之间的跳数。选择 H 范围内具有最高节点价值 ψ 的物理节点, 用于虚拟节点的重映射。虚拟节点重映射完成之后, 虚拟节点映射的变化会导致原来与该虚拟节点相关的虚拟链路不可用, 必须对这些虚拟链路也进行重映射, 才能完成服务承载网的愈合。先记录下与发生故障的物理节点存在虚拟链路映射关系的物理节点, 然后在选择完新的物理节点之后, 恢复新物理节点与那些被记录的物理节点之间的虚拟链路映射。可以采用直接重映射的方法进行一对一 (一条虚拟链路对一条物理路径) 的映射, 也可以采用多商品流的思想进行一对多 (一条虚拟链路对多条物理路径) 的映射。重映射方法的选择取决于网络中资源的数量。

假设有一条物理节点 N^p 发生故障, 有 k 个虚拟节点映射在 N^p 上, 记为 $N_1^v, N_2^v, \dots, N_k^v$ 。这 k 个虚拟节点分别包含在 r 个不同的服务承载网中, 记为 $V_1, V_2, \dots, V_r (r = k)$ ($r = k$ 是因为同一服务承载网的

虚拟节点必须映射到不同的物理节点)。物理节点 N^p 的故障会导致这 k 个虚拟节点的映射变得不可用, 同时这 k 个服务承载网也将不再正常工作。本文将进行 k 次恢复工作, 对于第 $i (1 \leq i \leq k)$ 次恢复, 愈合算法都会选择与 N^p 距离在 H 之内, 并具有最高 ψ 值的物理节点用来重映射虚拟节点 N_i^v , 并更新受到影响的虚拟链路, 来完成对应服务承载网 V_i 的愈合工作。在服务承载网 V_i 愈合之后, 都会更新网络中各个物理节点的 ψ 值, 来为下一个承载网 V_{i+1} 的愈合做好准备。

服务承载网 v_1 各虚拟链路的需求带宽以及底层网络 PN 的物理链路带宽如图 4 所示。为更方便地描述服务承载网的愈合过程, 图中同样不考虑物理节点的能力值 C_N^p 。物理节点 F 发生故障引发服务承载网 v_1 的虚拟节点映射 $M_1(b) = F$ 不可用。快速愈合算法在距离 F 固定为 2 跳的范围内 ($H = 2$) 选择节点 B 或节点 C 用作虚拟节点重映射。假设 2 个节点的可靠性分别为 $r_b = 0.5, r_c = 1$ 。底层网络中剩余的 4 条物理链路的可靠性均为 1。根据式(10)可知节点 B 和节点 C 价值的大小关系为 $\psi_b < \psi_c$ 。虚拟节点 b 映射到 C 之后 ($M_1(b) = C$), 愈合算法随之更新转移虚拟节点而发生变化的虚拟链路映射, 将 $M_1(ab) = ABF$ 更新为 $M_1(ab) = ABC$, 从而完成 v_1 的整个愈合过程。

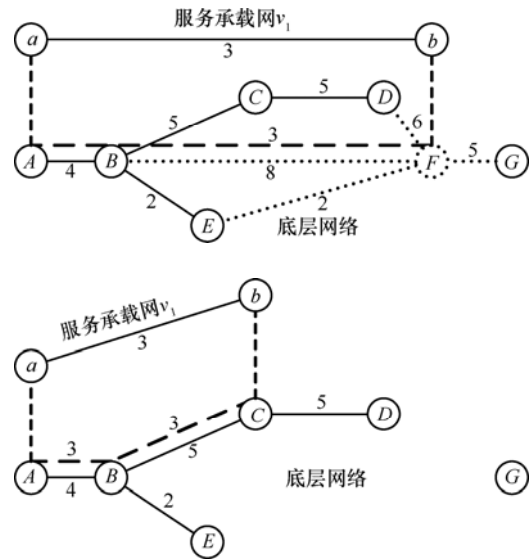


图 4 节点故障引发的服务承载网愈合

4.3 未能实时愈合的服务承载网的排队策略

在详细介绍完服务承载网的快速愈合算法之

后,还必须为那些受物理网络资源限制,未能及时愈合的服务承载网建立一套合理的排队调度策略,用于尽快处理那些服务承载网等待愈合的请求。利用一个优先级队列来存储承载网等待愈合的请求,这些请求按照到达的时间顺序被加入到队列中。该队列采用一种称为“可恢复优先(restorable first)”的调度策略来进行管理:底层网络的管理者会周期性地按序处理队列中的若干请求,一次周期内处理请求的数量取决于底层网络的处理能力。如果当前请求所指向的服务承载网可以愈合,则该请求会得到迅速响应,并在服务承载网愈合之后被移出队首;否则,该请求会被移动到队尾。这种调度策略可以第一时间判断当前的网络资源是否能满足队首的服务承载网愈合请求,从而在整体上降低了愈合请求的平均等待时间。

图 5 详细描述了“可恢复优先”调度策略的工作流程。假设在一个底层网络的处理周期 T 内可以处理 2 个服务承载网愈合请求,并且请求队列在初始状态存有 4 个请求: m_1 (队首), m_2 , m_3 和 m_4 , 分别代表第 1 到第 4 个请求。在第 1 个处理周期内, m_1 得到处理并被弹出队列,因其对应的服务承载网已成功愈合。 m_2 被移动到队尾,代表目前网络中的资源还不能支持其对应承载网的愈合工作;在第 2 个处理周期内, m_3 和 m_4 都被弹出队列,代表它们对应的承载网都可以成功愈合。并且又有一个新的请求 m_5 按照到达队列的时间先后顺序被加入到队尾,代表 m_5 对应承载网在愈合算法执行后未能成功恢复。

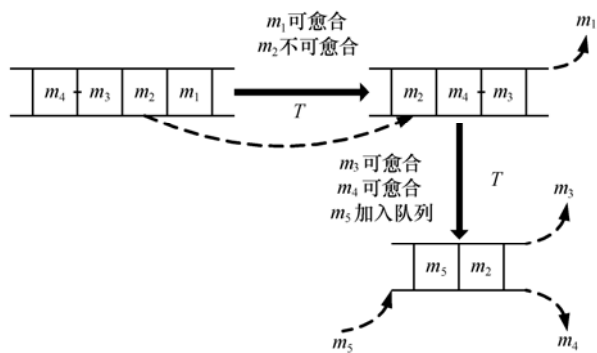


图 5 愈合请求队列调度

5 实验数据与性能评估

在本节中,本文通过 C++ 编程语言实现了服务承载网的快速愈合算法,并通过大量的仿真实验对

其在愈合成功率、愈合效率以及对底层网络负载均衡的影响 3 个方面进行了性能评估。本文还将服务承载网的快速愈合算法与传统的应对服务承载网故障时的全局重映射方法进行了对比,并通过实验数据说明本文提出的算法比传统的全局重映射算法在性能上有了一定的提高。

5.1 实验环境部署

1) 底层物理网络:底层物理网络的拓扑结构由 GT-ITM^[17] 拓扑生成工具生成:物理网络包含 50 个节点,每个节点之间的连接概率为 25%,即物理网络包含 300 条物理链路。物理节点的能力和物理链路的带宽在区间[15,20]内服从均匀分布。在进行仿真实验的过程中,物理节点故障次数和物理链路故障次数的比例为 1:4。这些故障均发生在事先预定的时间点。物理节点和链路的初始可靠性均为 1。
物理设备故障:在实验的过程中,预设了 50 次物理设备故障,包括 10 次物理节点故障和 40 次物理链路故障。这些故障会在每个固定的时间按顺序发生,并且发生过故障的节点和链路在实验过程中不会被修复。

2) 服务承载网:在仿真实验环境中,假设同时存在 50 个拥有随机拓扑结构的服务承载网运行在共享的底层网络之上。这些承载网的虚拟节点的能力需求以及虚拟链路的带宽需求在区间[5,10]内服从均匀分布。在仿真实验的过程中,通过服务承载网快速愈合算法来恢复因物理设备故障导致的承载网。

5.2 仿真实验数据

仿真实验比较了服务承载网快速愈合算法和传统的全局重映射方法在成功率,效率以及负载均衡 3 个方面的性能。从实验图表来看,仿真实验会在第 10 次、第 20 次、第 30 次、第 40 次以及第 50 次物理发生故障的时间点上对比两者在这 3 个方面的性能。

1)愈合成功率:服务承载网愈合的成功率 SR 定义如下,

$$SR = \frac{S_v}{F_v} \tag{11}$$

其中, F_v 代表失效的承载网个数, S_v 代表通过愈合算法成功恢复的承载网个数。如图 6 所示,快速愈合算法在整个服务承载网的愈合过程中保持了约 85% 的平均恢复成功率,而传统的全局重映射算法的平均成功率仅有大约 70%。特别是当物理网络故

障越来越多时，全局重映射算法的愈合成功率急剧下降。这是因为当物理网络中的资源紧缺时，快速愈合算法大量地重用了先前的网络映射，并允许虚拟链路进行一对多的映射，充分利用了网络中的剩余资源。而全局重映射算法只能进行虚拟链路的一对一映射，导致网络中一些剩余的带宽不能被充分利用，所以愈合成功率会大幅度下降。

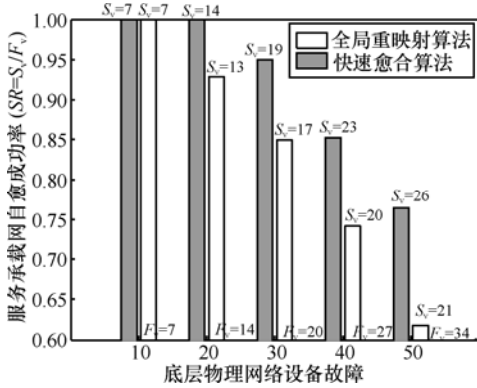


图 6 服务承载网愈合成功率

2) 愈合效率：在本文的仿真实验中，通过比较 2 种算法在愈合每个失效的服务承载网时所消耗的平均时间来进行算法的效率评估。从图 7 显示的实验结果可以看到，使用快速愈合算法进行愈合的服务承载网的平均愈合时间约为 0.5s 左右，而使用全局重映射算法进行愈合的服务承载网的平均愈合时间为 0.7s 左右，快速愈合算法的运行效率比全局重映射算法的效率提高了近 30%。这是因为快速愈合算法可以在一次运行的过程中愈合多个服务承载网，而全局重映射算法只能一次愈合一个服务承载网。

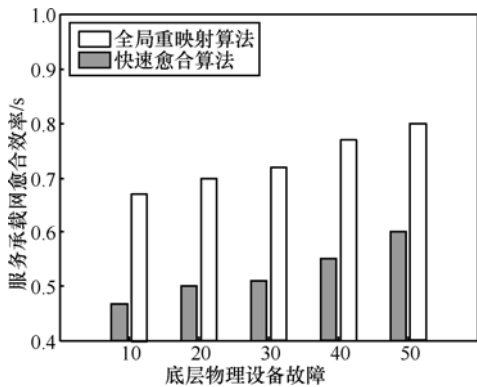


图 7 服务承载网愈合效率

3) 负载均衡：底层网络的负载均衡 Ω 包含物理节点的负载均衡 Ω_N 以及物理链路的负载均衡

Ω_L ，本文用标准差的形式来定义这 3 种负载均衡，定义如下：

$$\Omega_N = \sqrt{\frac{\sum_{i=1}^n (D_{Ni} - \overline{D_N})^2}{n}} \quad (12)$$

$$\Omega_L = \sqrt{\frac{\sum_{i=1}^l (D_{Li} - \overline{D_L})^2}{l}} \quad (13)$$

$$\Omega = \alpha\Omega_N + (1 - \alpha)\Omega_L \quad (14)$$

其中， D_{Ni} 表示物理节点 N_i 用于虚拟节点映射之后的剩余能力， $\overline{D_N}$ 代表底层网络中所有 n 个物理节点剩余能力的平均值；类似地， D_{Li} 表示物理链路 L_i 用于虚拟链路映射之后的剩余带宽， $\overline{D_L}$ 代表底层网络中所有 l 个物理链路剩余带宽的平均值。显然，根据标准差的特点，虚拟节点和虚拟链路分配在各个物理节点和物理链路之上的程度越是平均，底层网络的负载就越是均衡。对应地， $\Omega_N, \Omega_L, \Omega$ 三者的值就越趋近于 0。在仿真实验中，参数 α 的值由物理节点和物理链路发生故障的比例决定。因为节点故障和链路故障发生次数的比例为 1:4，所以 $\alpha = 20\%$ 。图 8 给出了在服务承载网愈合之后，底层网络中节点、链路以及全局负载的变化趋势。从图中可以看到 2 种算法在对物理节点负载的变化趋势方面的影响基本相同，这是因为 2 种算法所采用的虚拟节点重映射策略基本相同，都是选择一个适当的物理节点用作虚拟节点重映射。由于快速愈合算法允许虚拟链路进行一对多的映射，可以将虚拟链路相对平均地分配在各条物理链路之上，充分利用了网络中的剩余带宽，所以在服务承载网愈合之后，物理链路负载的增长趋势较为缓慢；而全局重映射算法只能进行一对一的链路映射，不能充分利用网络中的剩余带宽，所以在服务承载网愈合之后，物理链路负载的增长趋势比较明显。在本文的实验环境中，有 80% 的故障都来自物理链路，所以底层网络的全局负载主要受物理链路负载的影响。因此，两者的增长趋势也较为接近。

4) 网络规模变化：图 9~图 11 分别给出了当网络规模发生变化时，即底层网络物理节点和服务承载网的数量分别同时增加 10%、20%、30%、40%、50% 时（物理节点依然保持 25% 连接率），2 个算法

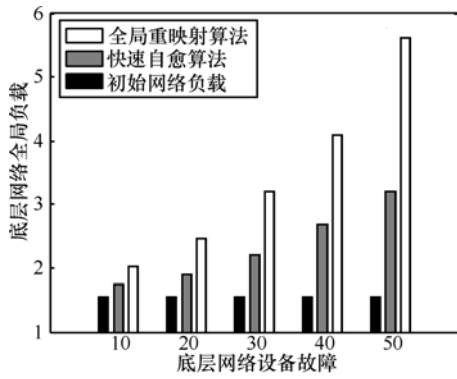


图 8 底层网络全局负载

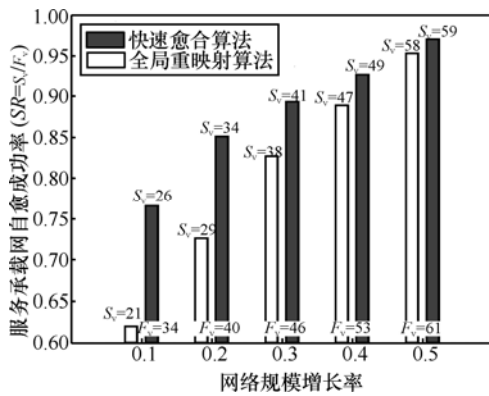


图 9 网络规模变化下的服务承载网愈合成功率

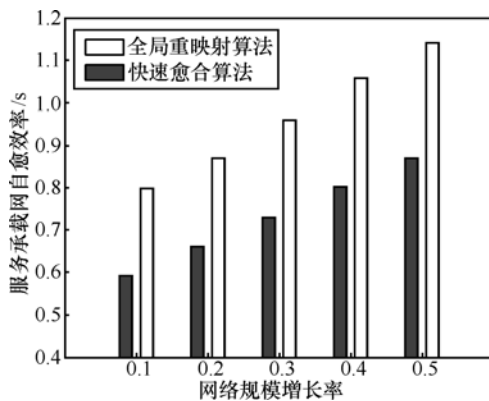


图 10 网络规模变化下的服务承载网愈合效率

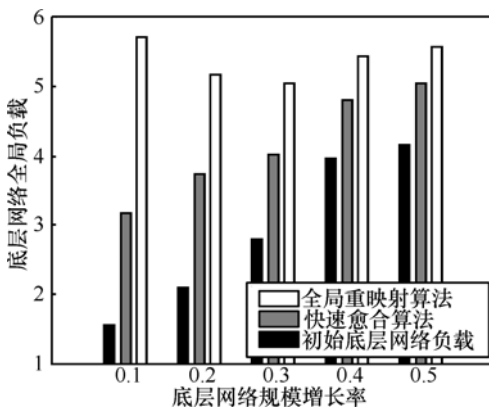


图 11 网络规模变化下的底层网络负载

在底层网络 50 次故障全部发生之后在愈合成功率,效率以及对底层网络全局负载 3 个方面的变化趋势。从实验结果可见,随着网络规模逐渐的增大,2 个算法在愈合成功率和负载均衡方面的差异越来越小。这是因为在大规模网络环境中底层网络资源丰富,当底层网络发生故障时,并不会导致网络资源的稀缺,快速愈合算法在网络资源稀缺时体现的性能优势相对弱化,所以在愈合成功率和负载均衡方面的差异变小;相反,当小规模网络发生故障时,很可能引起底层网络资源的稀缺,而快速愈合算法可以充分利用底层网络剩余资源,所以快速愈合算法在小规模网络环境中的性能更加突出。此外,服务承载网愈合效率,即愈合平均时间,是由算法的复杂度所决定。因为 2 个算法的复杂度相同,所以当网络的规模逐步从 10% 扩大到 50% 时,2 个算法在愈合效率方面的差异变化并不明显。

6 结束语

本文研究了可重构服务承载网的快速愈合机制,通过部分调整失效的虚拟节点或链路映射来愈合失效的服务承载网,避免了从全局上进行服务承载网重映射的繁琐工作。该方法基于网络流的思想,将服务承载网的愈合问题转化为多商品流问题加以解决,并通过解一次多商品流问题同时恢复所有失效的承载网,极大地提高了服务承载网愈合的成功率和效率,并且保持了底层网络的负载均衡。通过大量的仿真实验数据表明,该算法与传统的全局重映射方法相比,能更高效地提高服务承载网的愈合成功率和愈合效率,并可显著降低对底层网络负载的影响。

参考文献:

- [1] ANDERSON T, PETERSON L, SHENKER S, *et al.* Overcoming the Internet impasse through virtualization[J]. IEEE Computer Magazine, 2005, 38(4): 34-41.
- [2] GENI[EB/OL]. <http://www.geni.net/>.
- [3] BAVIER A, FEAMSTER N, HUANG M, *et al.* In VINI veritas: realistic and controlled network experimentation[A]. Proc ACM SIGCOMM[C]. Pisa, Italy, 2006.
- [4] AHUJA R K, MAGNANTI T L, ORILIN J B. Network Flows: Theory, Algorithms, and Applications[M]. Prentice Hall, 1993.
- [5] SZETO W, IRAQI Y, BOUTABA R, A multi-commodity flow based approach to virtual network resource allocation[A]. Proc of IEEE Global Telecommunications Conference (GLOBECOM'03)[C]. San Francisco, CA (USA), 2003. 3004-3008.

- [6] FAN J, AMMAR M. Dynamic topology configuration in service overlay networks—a study of reconfiguration policies[A]. Proc of IEEE INFOCOM06[C]. Barcelona, Spain, 2006. 1-12.
- [7] ZHU Y, AMMAR M. Algorithms for assigning substrate network resources to virtual network components[A]. Proc of IEEE INFOCOM06[C]. Barcelona, Spain, 2006. 1-12.
- [8] YU M, YI Y, REXFORD J, *et al.* Rethinking virtual network embedding: substrate support for path splitting and migration[A]. Proc of ACM SIGCOMM[C]. 2008. 17-29.
- [9] CHOWDHURY N M M K, RAHMAN M R, BOUTABA R. Virtual network embedding with coordinated node and link mapping[A]. Proc of IEEE INFOCOM09[C]. Rio de Janeiro, Brazil, 2009. 783-791.
- [10] LISCHKA J, KARL H. A virtual network mapping algorithm based on subgraph isomorphism detection[A]. Proc of ACM SIGCOMM[C]. Barcelona, Spain, 2009.
- [11] FAJJARI I, AIARI M, BRAHAM O. Towards an automatic piloting virtual network architecture[A]. Proc of IFIP International Conference on New Technologies, Mobility and Security (NTMS)[C]. Paris, France, 2011.
- [12] GUO T, WANG N, MOESSNER K. Shared backup network provision for virtual network embedding[A]. Proc of IEEE ICC2011[C]. Kyoto, Japan, 2011.
- [13] YALIAN P, XUESONG Q, SHUNLI Z. Fault diagnosis in network virtualization environment[A]. Proc of 18th Conference on Telecommunications (ICT)[C]. Ayia Napa, Cyprus, 2011.
- [14] AVRAMOPOULOS I, SYRIVELIS D, REXFORD J, *et al.* Secure Availability Monitoring Using Stealth Probes[R]. Princeton University, Technical Report TR-769-06, 2006.
- [15] FELDMANN A, GREENBERG A, LUND C, *et al.* Netscope: tra_c engineering for IP networks[J]. IEEE Network Magazine, 2000, 14(2):11-19.
- [16] AUGUSTIN B, CUVELLIER X, ORGOGOZO B, *et al.* Avoiding traceroute anomalies with paris traceroute[A]. Proc of Internet Measurement Conference[C]. New York, NY, USA: ACM Press, 2006. 153-158.
- [17] ZEGURA E W, CALVERT K L, BHATTACHARJEE S. How to model an internetwork[A]. Proc of IEEE INFOCOM[C]. Francisco, CA, USA, 1996.594-602.

作者简介：



缪宇霆（1986-），男，江苏南通人，浙江大学博士生，主要研究方向为网络虚拟化、可重构网络技术、P2P 网络等。



吴春明（1967-），男，浙江萧山人，博士，浙江大学教授、博士生导师，主要研究方向为人工智能、可重构网络技术、新一代高可信网络等。

杨强（1979-），男，河北正定人，博士，浙江大学讲师，主要研究方向为计算机网络，复杂网络与系统建模、控制、优化与仿真技术等。



姜明（1974-），男，江苏启东人，博士，杭州电子科技大学教授、硕士生导师，主要研究方向为计算机网络 QoS、网络虚拟化、可重构柔性网络、内容中心网络及网络多媒体技术等。