

网络拓扑发现综述

庄锁法, 龚 俭

(东南大学 计算机科学与工程系, 江苏 南京 210096)

摘要:网络拓扑发现是网络管理中一项非常重要的技术。文中从概述、主要的网络拓扑信息采集方法、网络拓扑发现算法三个方面对网络拓扑发现进行综述,在此基础上,设计出了一种针对 Internet 网络的拓扑发现算法和基于园区网的典型拓扑发现算法,对网络管理开发软件人员和网络管理人员具有较好的参考价值。

关键词:网络管理;拓扑信息;网络拓扑发现算法;SNMP

中图分类号:TP393.02

文献标识码:A

文章编号:1673-629X(2007)10-0080-04

A Survey on Network Topology Discovery

ZHUANG Suo-fa, GONG Jian

(Department of Computer Science & Engineering, Southeast University, Nanjing 210096, China)

Abstract: Network topology discovery is a key technology in network management. Give a survey from three points, including summarize, the collection technology of topology information and network topology discovery algorithm. On this base, design an algorithm of Internet topology discovery and a network topology discovery algorithm based on campus. Reaches a conclusion that this study is valuable to the network managers and the people who are engaged in the software development in network management.

Key words: network management; topology information; network topology discovery algorithm; SNMP

1 概述

网络管理的目的是通过监视和控制复杂的计算机网络,最大限度地保证其正常运行,并且要提高效率、降低成本。而网络拓扑发现则是配置管理的核心,故障和性能管理的基础,同时它也是衡量一个商业网络管理系统成败的重要尺度。因此拓扑发现算法的设计在整个网管系统的开发中有着举足轻重的地位。仅几年,国内外对拓扑发现算法的研究很多^[1~18],但都是基于某一局部范围或某一种协议的拓扑发现算法,所以不具有整体性。文中从概述、主要的网络拓扑信息采集方法、网络拓扑发现算法三个方面对网络拓扑发现进行综述。

1.1 基本概念

网络元素:指主机、路由器、交换机、网关、网桥、集线器等网络构成单元。

网络拓扑结构:指网络元素及它们之间的连接关系。

网络物理连接关系:指网络元素间有线或无线的

物理连接方式。这些连接方式有总线型、星型、环型、树型、星环型、混合型等。

网络逻辑连接关系:指网络元素间依照某种网络协议进行通信时所表现出来的逻辑上的连接关系。

网络拓扑发现:指利用计算机软件自动识别网络元素之间的物理或逻辑连接关系,确定网络的拓扑结构。

1.2 拓扑发现分类

(1) 按发现范围分为单管理域和多管理域。前者(参见文献[2~11])限于一个自治系统(Autonomous System, AS)内,如一个园区网;后者(参见文献[12~15])则跨越多个自治系统,如 Internet。

(2) 按发现协议层次常分为网络层和链路层。前者参见文献[1~4, 6~15])主要发现路由级 OSI 第三层联网设备路由器、子网及连接关系;后者(参见文献[5])重点发现交换式 OSI 第二层联网设备网桥、交换机等子网内设备详细连接关系。

(3) 按发现方法分为被动探测和主动探测。

1.3 网络拓扑发现算法的评价标准

实现拓扑发现的算法有很多种,但就性能和效率而言,判断起优劣的指标有:发现时延、网络负载、完整性、准确性。其中主要指标是网络负载、发现时延,具

收稿日期:2006-12-20

作者简介:庄锁法(1965-),男,江苏丹阳人,安徽科技学院副教授,主要研究方向为网络管理、网络安全;龚 俭,博士生导师,主要研究方向为网络管理、网络安全、网络行为学、新一代网络体系结构。

体地说:

发现时延:发现网络拓扑结构时算法需要运行的时间。

网络负载:算法运行时,占用网络和主机资源的多少。

完整性:算法能发现的网络元素及元素间连接关系占网络实际情况的百分比。

准确性:算法发现网络元素及元素间连接关系的准确率。

1.4 面临的困难

至今,网络的自动拓扑发现仍是一个富有挑战性的研究课题,研究出符合高效性、快速性、完整性及准确性四项目标的拓扑发现算法绝非易事。首先,网络本身没有提供专门针对自动拓扑发现的任何机制。拓扑发现工具的缺乏使得人们有时不得不采用网络的一些比较原始的工具,如 Ping 和 Traceroute 等。其次,不同的管理机构管辖着网络的不同部分。例如 Internet 就是由许多异构的自治系统(AS)组成,这些异构网络内部的硬件类型和运行软件都有可能不同。并且,不同的自治系统由不同的组织管理,它们有着各自的管理规则和安全策略,有的过滤甚至屏蔽了一些拓扑发现工具的搜索,这给网络的自动拓扑发现造成了一定的难度。

1.5 网络拓扑发现的研究范畴

(1)确定拓扑发现的规模。不同规模的网络,其互连设备、所需拓扑信息的详略、搜索开销的大小、发现速度的快慢等都会有所不同。因此,网络的拓扑发现应当针对不同规模(单管理域、Internet)网络,采用不同的协议工具设计算法。

(2)确定拓扑发现针对网络(OSI)哪一层以及采用的什么网络协议。基于第三层的网络拓扑发现要确定第三层设备的连接关系,即路由器与路由器之间、路由器与子网之间;基于第二层的网络拓扑发现要确定第二层设备的连接关系,即交换机、网桥和 hub 之间的连接关系。

(3)就网络拓扑信息的采集而言,应当确定是采用被动探测技术还是主动探测技术。主动探测技术是通过网管工作站主动向所有管理网络发送探测包,并采集返回的信息,进行分析最终形成网络拓扑。这种技术的优点是能够比较快地形成整个网络的拓扑,但是会给网络增加额外的流量。被动监测技术是通过在所有被观测的网络中都加入一个探测器,由它来采集信息,并发送到网管工作站形成拓扑。这种技术的优点是除了其本身向网管工作站递交拓扑信息以外,不产生额外的流量,网络负担小。但缺点是网络拓扑的形

成需要较长的时间。

(4)就拓扑信息的收集而言,应当确定采用何种工具收集信息,目前,信息收集的方式大致可以分为两种:一种采用简单网络管理协议 SNMP 来收集网络的拓扑信息;另一种信息收集方式采用比较通用的网络协议,如使用 DNS,RIP,OSPF,ARP 和 ICMP 等。

1.6 网络拓扑发现的应用

(1)模拟网络:为了模拟实际网络,必须首先得到该网络的拓扑结构;

(2)网络管理:涉及到配置管理、性能管理、故障管理,具体地说,网络拓扑信息可帮助网络管理者确定优化网络配置:测试和判断网络性能,发现网络瓶颈所在和失效的链路;

(3)服务器定位:网络拓扑信息可帮助用户确定自身在网络中的位置,从而确定服务器的位置,以及选择哪一个网络服务提供商可以将时延最小化、可用带宽最大化;

(4)为与拓扑结构相关的算法的性能改进提供依据;

(5)为 Internet 流量工程(traffic engineering)和网络行为学的研究提供基础辅助依据。

2 主要的网络拓扑信息采集技术

2.1 基于 SNMP 协议的拓扑信息采集方法

SNMP(simple network management protocol)是目前运行在 TCP/IP 上的网络管理协议,它是一个应用级协议,运行在 UDP 之上。SNMP 主要由三部分组成:管理信息结构,网络管理协议 SNMP 和管理信息库 MIB。其中 MIB 定义了可以通过网络协议进行访问的管理对象的集合,MIB-II(RFC1213)是被设计用描述所有网络元素(主机、路由器和网桥等)所必需的重要信息。在 MIB-II 中,与拓扑相关的关键信息有:IfTable 包含路由器的接口信息,见表 1;IfRouteTable 包含路由器中的路由表,见表 2;IfAddrTable 包含了路由器的地址信息,见表 3。

表 1 iftable 表

ifindex	ifdescr	iftype
接口索引号	接口描述	接口类型

表 2 iproutetable

IpRouteDest	IpRouteIfIndex	IpRouteType	IpRouteNextHop
目的 IP 地址	本地接口的索引值	路由类型	下一站的地址

表 3 IfAddrTable

IpAdEntAddr	IpAdEntIfIndex	IpAdEntNetMask
实体的 IP 地址	接口的索引值	与该实体的 IP 地址相关的子网掩码

用户利用报文 GetRequest-PDU 或 GetNextRe-

quest-PDU 与目标代理交互,从目标代理进程的 MIB-II 中提取网络设备的配置信息,通过分析算法,形成拓扑图^[16]。

2.2 基于 ARP 协议的拓扑信息采集方法

任何有以太网接口的网络设备都必须支持地址解析协议(ARP, Address Resolution Protocol),并在本机维护着一张 ARP 表,用于 IP 地址与以太网地址间的地址解析和转换。

读取节点系统上的 ARP 高速缓存 cache 的 ARP 表,可以发现与其各以太网端口相连的局域网中的所有设备,再判断网络中的路由器和交换机,并继续根据 ARP 表进行发现,从而得到整个以太网的拓扑结构。

2.3 基于 ICMP 协议的拓扑信息采集方法

ICMP 协议作为 IP 协议的一部分,用户通过向给定地址空间内所有可能的主机发送 ICMP“回显请求”包来确定活动主机,记录其 IP 地址,子网掩码,利用 traceroute 来发现给定的主机路由信息,根据这些信息,通过拓扑发现算法程序分析,从而得到拓扑图。

2.4 基于 RIP 协议的拓扑信息采集方法

RIP 是一种距离向量路由协议,要求路由器每个节点存放到达各个目标站点的距离,所谓距离即指到达目标节点所经过的跳数。RIP 协议要求每隔一定的时间(例如 30s),各路由器向其相邻路由器发送自己存放的、到达各目标主机的距离信息,即广播自己的路由表的所有路由表项。同时接收其它相邻路由器发送来的路由表项更新报文,将其距离值与其本身的路由表中的信息进行比较,不断根据这些信息来更新路由表。用户从路由表中得到所需要的 IP 地址和路由器的邻接关系。

2.5 基于 DNS 的拓扑信息采集方法

DNS 系统主要用于网络设备 IP 地址到名字的映射,同时也维护一些其他信息,例如电子邮件的选路信息,设备的硬件平台及操作系统。DNS 模块的主要任务是向 DNS 服务器发送区域传送命令(Zone transfer),通过从某一结点开始递归下降遍历 DNS 树而返回该区域内的所有域名列表,因此可以用来发现区域内的所有主机和路由器。

2.6 基于 OSPF 协议的拓扑信息采集方法

OSPF 路由协议^[17]是开放最短通路优先协议(Open Shorted Path First),也属于内部网关协议,其基本思想非常简单,每个结点存放网络的链路状态图,路由器之间通过相互交换路由信息来更新链路状态图。利用 OSPF 路由信息协议可以从路由设备中提取通过该设备的所有 IP 地址域、IP 子网的路由信息表,通过相关的拓扑算法,构造出拓扑连接信息表,从而得到拓

扑图。

以上每一种信息采集方法对应的拓扑发现算法,它们的性能优劣涉及到算法的收敛速度、网络开销、拓扑的完整性、算法实现的难易程度,分析比较见表 4。

表 4 拓扑发现算法比较表

方案	路由重复	收敛速度	拓扑精度	网络负载	算法难易	应用范围	网络规模
SNMP	中	慢	中	低	简单	大	大
ARP	低	快	高	中	简单	中	中
ICMP	高	中	高	高	简单	小	小
RIP	高	中	高	高	难	小	中
DNS	高	中	中	高	难	中	小
OSPF	高	慢	低	高	难	小	小

由上面的分析可以得到以下结论:

(1)每一种方法都有一定的优势,但也或多或少存在一些不足。每种方法有其适用的范围,所以要想对于各种网络的各种设备进行搜索,应该根据实际需要有针对性地结合它们。

(2)大型网络如 Internet 网络拓扑发现可以利用 OSPF, RIP, BGP 等路由协议获得路由信息,得到相应的 IP 地址,再用 Ping, Traceroute 工具探测目的主机的可达性和路由器的连接信息。

(3)一个管理域内的网络拓扑发现可利用 SNMP 协议发现主干网络拓扑结构,再利用 ARP, ICMP 协议发现各子网内部的拓扑结构。

3 网络拓扑发现算法

3.1 Internet 网络拓扑发现算法

Internet 规模巨大,结构复杂,本质上具有动态性,同时流量也具有突发性以及人们对 Internet 拓扑结构本身了解甚少,因此,局域网的方法和技术不适合于广域网,在 Internet 拓扑发现研究中人们不能作较多的前提假设,只能依靠探测手段。目前进行大型网络拓扑发现主要是利用所有路由器都支持的 ICMP, UDP 协议,进行主动探测^[12~15]。

算法的主要思想:在 IP 网络的任一节点主动地探测网络拓扑,根据前一次的探测结果和 IP 地址分配规律,利用启发式方法确定下一次的地址探测目标,提高算法效率。

Internet 网络中 IP 地址分配规律:

规则 1:利用优化发现路径 {1, 129, 254, 2, 65, 533, 193, 11, 17, 101, 127} 快速推断当前网络是否活动。

规则 2:利用发现的活跃 IP 地址,猜测子网掩码,推断当前网络中的其他活动 IP 地址。

规则 3:利用已知的网络前缀,根据 CIDR 聚合^[18]

地址原理,猜测网络前缀长度,推断其它活动网络前缀。

算法描述:

(1)根据 RIP,OSPF,BGP 等路由协议,获得初始 IP 地址组,称为临时 IP 地址组;

(2)利用 PING 工具探测临时 IP 地址组中的每一个临时 IP 地址的活动性,如果可达,将该 IP 地址存放活动 IP 地址组;

(3)利用规则 1~3 推测出可能的活动 IP 地址不重复存入临时 IP 地址组;

(4)利用 TRACEROUTE 工具,对活动 IP 地址,探测路由器的连接信息;

(5)利用分析程序,构造路由器间的连接关系,路由器和子网的连接关系。

3.2 园区网拓扑发现算法

园区网络拓扑发现算法的研究主要分为三类:基于 SNMP 协议的拓扑发现算法(典型的可参见文献[1,4,6]);基于 OSPF 路由协议的拓扑发现算法(见文献[7~9]);综合运用各种协议的拓扑发现算法(见文献[10,11])。

拓扑图尤其是连接关系的合理定义对拓扑发现、数据结构设计和拓扑显示意义重大,所以采用如下定义:以 $G(V, E)$ 表示拓扑图,以 V 表示结点(路由器), E 为连接,定义为 (N, R) 表示路由器与子网相连,确切地说是路由器的某个接口属于子网 N 。算法采用分级发现的策略,将网络拓扑发现分成两级进行:第一级采用基于 SNMP 协议的拓扑信息采集方法发现路由设备和子网以及连接关系;第二级利用 ARP、ICMP 协议发现子网中的第二层设备和主机设备以及连接关系。

园区网拓扑发现典型算法描述:

(1)初始化路由器集合为空,加入搜索起点路由器。

(2) while(路由器集合不为空)

{

取出一个路由器,

·读地址表。根据第 2.1 节 IP 地址表所述,将地址表每行的 IpAdEntAddr 和 IpAdEntNetMask 按位与,得出路由器在相应接口上连接的子网地址,即为连接 (N, R) 。

·读路由表。根据第 2.1 节 IP 路由表所述,将路由类型 IpRouteType 为 4(indirect)所对应的下一站路由器 IpRouteNextHop 作为新发现的路由器,不重复地加入路由器集合中。

}

(3)子网搜索:根据(2)可获得路由器的连接的所有子网地址,结合 ARP 和 ICMP,根据 ARP 协议原理,先用 ARP 表确定子网中实际连接的主机以及主机的 IP 地址,再 PING 这些实际的 IP 地址确定主机的活动性,这种方法极大地降低了传统 PING 方法搜索子网耗费的大量时间,而且不会增加网络负载。

4 展望与总结

拓扑信息收集是网络拓扑发现的核心工作,提高自动收集方法不是唯一途径,还需要有更多组织与个人共同参与,提供丰富全面的信息。P2P 是一种基于互联网环境的新的应用型技术,如何充分发挥 P2P 网络的功能,研究 P2P 网络拓扑结构,是一个非常意义的课题。关于 IPv6 网络拓扑发现算法的研究,一直是人们研究的热点。网络拓扑发现是网络管理中一项非常重要的技术。文中的网络拓扑信息采集方法、网络拓扑发现算法对网络管理开发软件人员和网络管理人员具有较好的参考价值。

参考文献:

- [1] 李可,薛质,铁玲. IP 网络拓扑自动发现研究[J]. 计算机工程,2004(3):66-68.
- [2] 瞿宏阳. 网络拓扑顺序搜索算法及其 Web 实现[J]. 微电子学与计算机,2001(3):41-45.
- [3] 瞿宏阳. 层次式网络拓扑显示方法的设计与 Web 实现[J]. 微电子学与计算机,2002(4):60-62.
- [4] 卞正皓,王汝传. 基于 SNMP 网络拓扑图构造方法的研究[J]. 南京邮电学院学报,2003,23(2):74-77.
- [5] 郑海,张国清. 物理网络拓扑发现算法的研究[J]. 计算机研究与发展,2002,39(3):264-268.
- [6] 廖丽惠,陈琳,龚正虎. 网络管理系统中的拓扑发现[J]. 计算机工程与应用,2003(30):172-174.
- [7] 徐大海,黄健强. 基于 OSPF 的网络拓扑搜索[J]. 计算机工程与科学,1999,21(6):17-21.
- [8] Shaikh A, Goyal M. An OSPF Topology Server: Design and Evaluation[J]. IEEE Journal on Selected Areas in Communications,2002,20(4):746-755.
- [9] 徐建锋,邓永平,丁圣勇. 基于 OSPF 服务器的网络拓扑发现[J]. 计算机应用,2004(8):98-100.
- [10] 邱建林,何鹏. 一种改进的网络拓扑发现方法[J]. 计算机应用,2005(4):891-893.
- [11] 王志刚,王汝传,王绍棣,等. 网络拓扑发现算法的研究[J]. 通信学报,2004(8):36-43.
- [12] Siamwalla R, Sharma R, Keshav S. Discovering Internet Topology[C]//Proc. of IEEE INFOCOM. [s. l.]:[s. n.], 1999.
- [13] 林宏刚,李焕洲,戴宗坤. 一种启发式拓扑发现算法[J]. 四

(下转第 91 页)

微大一些的块。动态调度的灵活性最大,但是在调度出错时,造成的性能损失也最大。对于平衡良好的代码,应避免使用非常小的块。本例子中调度块大小为 128 明显是很好的选择。而对于那些不平衡的代码,正确的块大小应该在运行较少的块和达到更佳的线程平衡之间做一个权衡。引导调度使用较小的块作为限制时运行最佳、灵活性最大。限定为较大的块大小时,运行的时间将很长。引导调度的上限是循环总大小的二分之一,对于平衡良好的代码,应该相当于在一次运行中将循环拆分为相等几部分的静态情况。

3 结 语

OpenMP 是针对共享地址空间并行计算机提供的并行计算库,目前 Microsoft 和 Borland 公司都有相应产品支持 OpenMP2.5。使用 OpenMP 不必写诸如 CreateThread 之类的线程管理代码,编写多线程程序简便高效,而且 OpenMP 提供了丰富的指令,对于同步共享变量、合理分配负载等任务,都提供了有效的支持。不过 OpenMP 也存在着一些不可避免的缺点:第一,OpenMP 主要以预编译指令(#pragma)实现多线程并行,所以在单核机器上编译的程序在多核机器上运行时无法体现多核的优势;第二,OpenMP 对编译器要求比较高,一般要求 Microsoft Visual Studio 2005 或者需要 Intel 编译器。不过长远来讲,OpenMP 的优势是明显的。Intel 技术官曾说“今后的处理器发展是内部优化与集成多核而不是单纯地提升处理器的频率,采用多线程的软件也将会是今后软件的主流。”充分发挥多核处理器的优势使软件性能最优化,给中国带来

一个极大的机遇。开发 OpenMP 软构件库需要大量的人力资源,这项工作适合在中国进行。OpenMP 库在未来的经济效益可以同微软的操作系统相比,后者为用户顺利使用 PC 提供工具软件支持,前者为顺利使用多处理器提供库函数支持。世界正在进入多处理器时代,OpenMP 库将成为程序员必不可少的工具。

参考文献:

- [1] 陈国良.并行算法实践[M].北京:高等教育出版社,2004.
- [2] Quinn M J. Parallel programming in C with MPI and OpenMP [M].北京:清华大学出版社,2005.
- [3] 赖建新,胡长军,赵宇迪,等. OpenMP 任务调度开销及负载均衡分析[J]. 计算机工程,2006,18:(sup)58-60.
- [4] Grama, Ananth. Introduction to parallel computing [M]. 北京:机械工业出版社,2003.
- [5] Foster I, Designing and building parallel programs [M]. 北京:机械工业出版社,2002.
- [6] Andrew. Multithreading parallel and distributed programming [M]. 北京:高等教育出版社,2002.
- [7] Malyszhkin V. Parallel computing technologies [C]//8th international conference, PaCT 2005. Krasnoyarsk, Russia, 2005. Berlin; New York: Springer, 2005.
- [8] 杨淑莹. VC++ 图像处理程序设计 [M]. 北京:清华大学出版社,2003.
- [9] Dongarra J. Parallel computing programming [M]. 北京:电子工业出版社,2005.
- [10] Wilkinson B, Allen M. Techniques and applications using networked workstations and parallel computers [M]. 北京:机械工业出版社,2005.
- [11] 川大学学报:工程科学版,2005,37(3):118-122.
- [12] Govindan R, Tangmunarunkit H. Heuristics for internet map discovery [C]//Proceedings IEEE INFOCOM. [s. l.]: [s. n.], 2000:1371-1380.
- [13] 姜 誉,胡铭曾,方滨兴,等. 一个 Internet 路由器级拓扑自动发现系统[J]. 通信学报,2002,23(2):54-62.
- [14] 杨家海,任宪坤,王沛瑜. 网络管理原理与实现 [M]. 北京:清华大学出版社,2000.
- [15] Peterson L L, Davie B S. Computer networks: A Systems Approach [M]. 3rd Edition. USA: Elsevier Science, 2003.
- [16] Rekhter Y. An architecture for IP address allocation with CIDR [S]. RFC1518, 1993.
- [17] Maedche A, Neumann G, Staab S. Bootstrapping an Ontology-based Information Extraction System [C]//Intelligent Exploration of the Web, Studies in Fuzziness and Soft Computing. Heidelberg: Physica-Verlag GmbH, 2003:345-359.
- [18] Yeh C L, Su Y C. Web Information Extraction for the Creation of Metadata in Semantic Web [C]//Proceedings of ROCLING. Tainan, Taiwan: [s. n.], 2005.
- [19] erogeneous, Distributed Biological Data Sources [C/OL]//Proceedings of the IJCAI2001 Workshop on Knowledge Discovery from Heterogeneous, Distributed, Autonomous, Dynamic Data and Knowledge Sources. 2001. http://www.cs.iastate.edu/honavar/Papers/ijcaiworkshoppaper.pdf.
- [20] Vargas-Vera M. Knowledge Extraction Using an Ontology-Based Annotation Tool [C]//Workshop on Knowledge Markup & Semantic Annotation. [s. l.]: ACM Press, 2001:5-12.

(上接第 83 页)

(上接第 86 页)