

# 综合布线测试技术在网络流量分析中的应用

刘广聪,王 康,张 桦  
(广东工业大学,广东 广州 510006)

**摘 要:**综合布线测试技术是用电缆、光缆测试仪对综合布线工程进行的现场验收测试的技术。数据仓库技术是分析历史的、海量数据的热门技术。通过介绍数据仓库技术以及综合布线技术的有关概念,并将两种技术运用到网络流量分析中去。利用了数据仓库的聚合技术,通过确定不同粒度的维度表,实现了对网络流量的海量数据的高效处理。通过该方法可以进行网络路由调整并且能够通过对现有数据的分析,预测未来可能发生的故障。提高了对网络流量的分析能力和决策水平。

**关键词:**数据仓库;OLAP;网络综合布线测试;网络流量分析

**中图分类号:**TP393.08

**文献标识码:**A

**文章编号:**1673-629X(2011)03-0239-04

## Application of Cabling Testing Technology in Network Traffic Analysis

LIU Guang-cong, WANG Kang, ZHANG Hua  
(Guangdong University of Technology, Guangzhou 510006, China)

**Abstract:** Cabling testing technology is a technology which uses network tester testing network cabling system. A data warehouse is a repository of an organization's electronically stored data. Data warehouses are designed to facilitate reporting and analysis. Introduce the network cabling test technology and data warehouse technology and use them to analyze network traffic. Utilize polymerization of data warehouse to realize efficient data processing for network traffic. The adoption of this method can help us adjust routing, and forecast the network failure. Improve the capacity of the network traffic analysis and decision-making.

**Key words:** data warehouse; OLAP; cabling testing technology; network traffic analysis

## 0 引言

随着互联网技术的日益普及,网络的应用也越来越广泛。网络的大规模使用也对网络的管理和维护工作提出了较高的要求。网络管理工作中,网络性能管理是网络管理中的重要部分。在现实的应用中,大规模局域网内往往由于网络设备使用不当或者使用期限过长等因素导致网络性能急剧下降。针对这种物理层故障导致的网络性能的下降问题,可以将数据仓库及布线测试技术引入到网络流量分析中,可以起到一定的故障预警作用。

## 1 网络综合布线测试技术

### 1.1 验证测试和认证测试

布线测试一般分为两个部分,电缆的验证测试和电缆的认证测试。

布线的验证测试主要是测试线路的基本安装情况。如双绞线有无开路、短路,线序是否正确连接,双绞线走向如何等等。双绞线的一个特殊故障是串扰。所谓串扰就是将原来的两对线的双绞结构打开了。或者是并没有按照标准线序进行打线。这种故障的端到端的连通性是好的,所以用万用表是查不出来的,只有用专用的线缆测试仪器才能检查出来。串扰故障不易发现,这是由于当网络在低速运行或流量很低的情况下,这种故障的表现不是很明显。单在网络高速运行时,串扰故障对网络有极大的影响。

所谓的布线认证测试是指线路除了正确的连接以外,还要满足相关的标准。认证是测试网络布线的最高层次技术,它具有认证线缆传输数据效率的能力。认证测试将确保网络的性能并保证网络可以始终运行在最佳状态<sup>[1]</sup>。即安装好的电缆的电气参数,包括衰减、串扰、回波损耗等电气参数是否达到有关标准。只有通过认证测试,才能保证线路在高速运行情况下的稳定、通畅。

关于综合布线的标准,国际上主要有两大标准,

收稿日期:2010-06-18;修回日期:2010-09-24

基金项目:2009 广东省自然科学基金项目(090147)

作者简介:刘广聪(1970-),男,高级工程师,研究方向为计算机网络与分布式数据处理、网络安全,数据挖掘。

TIA 和 ISO。TIA 制定美洲的标准, TIA 对 ISO 标准也有影响。我国线缆来源主要是美国, 因此我国的布线标准更多使用 TIA 标准。

## 1.2 认证测试内容

针对现在普遍使用的双绞线, 综合布线认证测试对双绞线的主要测试内容有接线图、长度、传输时延和时延偏离、插入损耗、近端串扰和综合近端串扰、衰减串扰比和综合衰减串扰比、回波损耗、等效远端串扰和综合等效远端串扰等。下面简要介绍一下插入损耗、近端串扰和综合近端串扰、衰减串扰比和综合衰减串扰比、回波损耗、等效远端串扰和综合等效远端串扰。

**插入损耗:**插入损耗是指发射机与接收机之间, 插入电缆或元件产生的信号损耗, 通常指衰减<sup>[2]</sup>。

**近端串扰和综合近端串扰 (Near End Cross Talk, 缩写 NEXT):**也叫线对-线对 NEXT 损耗 (Pair-Pair Next Loss), 是指在与发送端处于同一边的接收端处所感应到的从发送线对感应过来的串扰信号。在串扰信号过大时, 接收器将无法判别信号是远端传送来的微弱信号还是串扰杂讯。综合近端串扰 (Power Sum NEXT, 缩写 PSNEXT) 表示的是双绞线在同一时间一个线对受到另外 3 个线对的影响<sup>[2]</sup>。

NEXT 的计算公式如下:

$$\text{NEXT} = -20 \times \log(10^{-\frac{\text{NEXT1}}{10}} + 2 \times 10^{-\frac{\text{NEXT2}}{10}}) \text{ dB}$$

其中 NEXT1 是水平电缆的近端串扰损耗, NEXT2 是连接构件的近端串扰损耗。高的 NEXT 值反应了小的近串扰值和更好的电缆性能。在所有的网络运行特性中, 串扰值对网络的性能影响是最大的。

PSNEXT 的计算公式如下:

$$\text{PSNEXT} = -10 \times \log(10^{-\frac{\text{NEXT1}}{10}} + 10^{-\frac{\text{NEXT2}}{10}} + 10^{-\frac{\text{NEXT3}}{10}}) \text{ dB}$$

式中 NEXT1, NEXT2 和 NEXT3 是选定线对与其他 3 线对之间对与对串扰测量值。

**衰减串扰比和综合衰减串扰比:**衰减串扰比 (ACR) 表示的是链路中有效信号与噪声的比值。衰减串扰比的定义为: 在受相邻发信线对串扰的线对上其串扰损耗 (NEXT) 与本线对传输信号衰减值的差值, 单位为 dB。综合衰减串扰比 (PSACR) 是根据 PSNEXT 和插入损耗之间的差别算出的值, 它是信噪比的一种, 单位为 dB。PSACR 反映了三对线同时进行信号传输时对另一对线所造成的综合影响。它只要用于保证布线系统的高速数据传输, 即多线对传输协议<sup>[2]</sup>。

双绞线中每个线对的 ACR 和 PSACR 的计算公式如下:

$$\text{ACR} = \text{NEXT} - A \text{ (A 为衰减)}$$

$$\text{PSACR} = \text{PSNEXT} - A \text{ (A 为衰减)}$$

**回波损耗:**由于电缆的结构无法完全一致, 因此会

引起阻抗发生少量变化。阻抗的变化会使信号在电缆中传输时被反射回来从而产生损耗。信号在电缆中传输时被反射回来的信号能量强度就是回波损耗 (Return Loss, 缩写 RL)。它是以分贝 (dB) 为单位表示的。回波损耗对于全双工传输的应用非常重要。

**等效远端串扰和综合等效远端串扰:**远端串扰 (FEXT) 是指信号在接近电缆远端处所感应到的从发送线对感应过来的串扰信号。因为串扰信号经过衰减到达远端已经很小了, 所以远端串扰的影响不会很大, 因此远端串扰 (FEXT) 不会作为一个单独的测试项目。

而等效远端串扰 (ELFEXT) 就是为补偿 FEXT 测量值的衰减量。直接从 FEXT 值减去衰减值, 即可得到等效远端串扰。

$$\text{ELFEXT} = \text{FEXT} - A \text{ (A 为衰减)}$$

**综合等效远端串扰 (PSELFEXT) 的“综合”**是指远端串扰的综合, 它是指某线对受其他线对综合的等效远端串扰影响, 用分贝 (dB) 表示。这里, 干扰源不再是等效远端串扰测试中的单一线对, 而是 3 对线共同的干扰。

在人们的日常使用过程中, 有很多因素会影响到以上各个参数的测量结果。比如双绞线被重物挤压变形。这就破坏了原来双绞线的绞合结构, 极易发生近端串扰故障。双绞线在铺设过程中受力不均, 导致双绞线的阻抗发生变化, 容易引发回波损耗的异常。以上故障, 表现形式往往为感觉网速变慢、数据传输不畅、丢包率增加等, 不易引起用户的注意, 但却严重影响了网络的正常使用。

## 2 数据仓库及布线测试技术在网络流量分析中的应用

数据仓库技术是分析历史的、海量数据的热门技术。数据仓库是一个面向主题的、集成的、时变的和不易失的数据集合, 支持管理部门的决策过程<sup>[3]</sup>。数据仓库中的数据来源一般都是日常事务性数据库中的数据。数据库中的数据经过抽取、清理等过程进入数据仓库中。等待用户进行数据分析。关于数据仓库的研究是信息领域的前沿课题, 它的兴起也给信息社会带来了新的契机, 构建成功的数据仓库已经成为许多行业的迫切需要, 在原有应用系统上建立数据仓库, 并以联机分析处理和数据挖掘为手段进行数据分析已成为该领域研究的热点<sup>[4-8]</sup>。从数据仓库技术的特点来看, 非常适合网络流量分析的工作。而多维数据分析是数据仓库系统中最常用的分析技术之一。它又被称为联机分析处理 (Online Analytical Processing, OLAP)<sup>[9]</sup>。通常来说, 建立网络流量分析的数据仓库的目的是为了通过多方面的汇总信息, 来发现网络资

源的利用率和IP地址与流量增长之间的一定规律,用以完善网络资源的合理利用和异常情况的监控。之前也有许多学者做过类似的研究工作。但很少有人将网络布线的最底层,即物理层的因素考虑在内。但是在实际的应用中,网络物理层的一些故障,也是影响网络流量的重要因素。在网络流量数据仓库中,建立相关维度可以提高网络流量的分析的全面性。

2.1 网络流量分析数据库的建立

网络流量分析数据库用于记录日常网络流量信息。但单纯的网络流量分析数据库只能提供给用户简单的增删查改的功能。提供不了分析功能。因此建立了网络流量分析数据库后,要将数据库中的数据经过ETL汇总到数据仓库。

针对网络流量分析的这个目标,建立如下数据表(见表1),进行流量数据搜集。

表1 字段定义

字段名	类型	字段功能描述
SRC_IP	字符串	存放数据包源地址
DEST_IP	字符串	存放数据包目的地址
SRC_SUBNET	字符串	源地址所属子网
DEST_SUBNET	字符串	目的地址所属子网
IP_SIZE	整形	数据包大小
IP_TIME	日期	数据包传送时间
PSNEXT	浮点型	综合近端串扰余量
PSACR	浮点型	综合衰减串扰比余量
PSFEXT	浮点型	综合等效远端串扰余量
RL	浮点型	回波损耗余量

上表中除了一些基本的数据包的信息如原目的IP地址以及端口号、数据包大小、传送时间等基本信息外,还加入了一些布线的测试结果。PSNEXT、PSACR、PSFEXT,以及RL的测试结果也会随着使用年限以及周围环境的变化产生变化。比如温度、辐射、网线的挤压变形、损耗等。这种变化有时并不会导致网络出现断网等情况。但是可能会导致网速骤减,数据包丢失率增加等情况。也会严重影响网络的使用。不过由于网络测试结果的变化频率不会很快,所以可以制定一个相对较大的时间频率进行测试。比如每个月进行一次认证测试,并记录相应的数据,录入表中。

2.2 数据仓库建模

首先确定维表和事实表。主题涉及的维表有源地址维表、目的地址维表、时间维表、PSNEXT结果维表、PSACR结果维表、PSFEXT结果维表、RL维表。事实表为流量表。维表和事实表之间采用星型模型(如图1所示)。

维度粒度的确立是数据仓库建模的非常重要的一个方面。多维分析的一个显著优点是可以对维度进行

下钻和上钻操作。适当地确立维度粒度也很重要。合适的粒度划分便于生成概念层次树和从概念层次上进行分析<sup>[10~12]</sup>。对源IP地址和目的IP地址可以按“子网-IP地址”进行划分。时间维可以按照“分-时-日-月-年”进行划分。

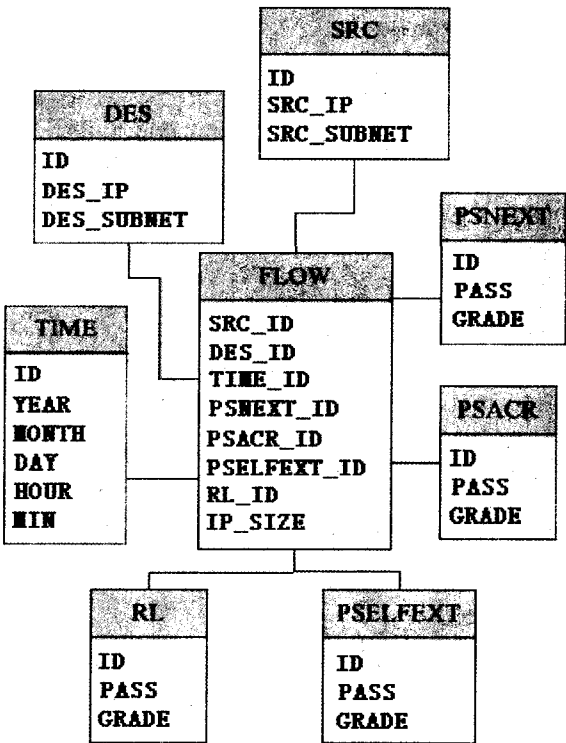


图1 星型模型

PSELFEXT维,PSNEXT维,PSACR维,RL维都属于测试结果维。当使用网络布线测试仪器进行测试的时候,测试结果是分布于0~100dB的曲线(见图2),

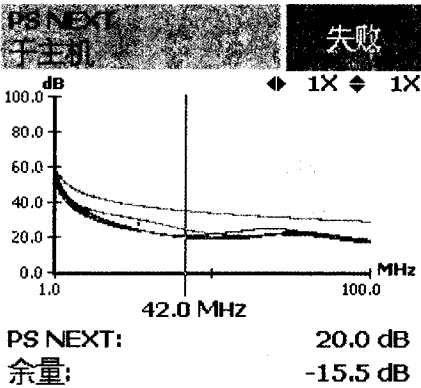


图2 测试结果

并且每个频率的值都不一样,所以测试结果值不适合直接作为评价网络性能好坏的标准。但是可以使用测试仪器提供的余量值作为标准。余量表示的是当前测试结果与测试标准的差距。如果余量为正,测试结果属于标准数据范围以内并且绝对值越大,性能越好。测试结果为负则不达标,绝对值越大,越不符合标准。

所以可以对测试出来的余量进行归约。为其设置一个等级 (GRADE)。规约表如表 2 所示。

表 2 规约表

余量值范围	等级
余量 < -10dB	1
-10dB < 余量 < 0dB	2
0dB < 余量 < 10dB	3
10dB < 余量	4

因此 PSELFEXT 维, PSNEXT 维, PSACR 维, RL 维可以按照 (等级-是否通过测试) 进行划分。维度粒度确立后, 星型模型已经建立完成。可以通过对各个维度进行上钻和下钻操作, IP 数据包进行分析。可以获得指定源 IP 到指定目的 IP 的某段时间内的, 网络测试结果等级均为 1 的 IP 数据包的总量。如果网络测试结果等级为 1 的 IP 数据包总量比较大, 就必须做出应对, 及时检查设备、网线, 做出修理或更换。

### 3 数据仓库建模及实验结果分析

SQL2000 提供了一个数据仓库建模组件“Analysis Services”。利用此工具可以进行数据仓库建模, 并能查看多维数据模型并且进行 OLAP 分析。详细建模过程在此不做赘述。

模型简历成功后, 单击维度选项中的加号或者减号就可以进行多维数据集的上钻或下钻操作, 比如用户可以查看时间为 2000 年第一季度 1 月份从源 IP 为 192.168.0.2 发出的各项网络测试等级为 1 的所有累计数据包的数量为 4454589。并可按季度, 按年进行聚合查看。聚合的粒度越粗, 压缩效果越好, 细节数据也会越少。并且用户可以在时间维度或其他维度进行聚合查看。通过等级为 1 的数据包的数量和总的数据包数量的比值, 可以对网络的健康状况进行量化的评估。或者可以单纯查看某段时间内每个源 IP 及目的 IP 的流量情况。

选择合适的维度条件进行 IP 数据包累计流量的查看, 可以帮助用户进行实时动态路由调整, 以及故障

预警。

### 4 结束语

将布线测试技术引入到流量分析中, 是一个新的尝试, 它的应用可以协助网络管理员更好地管理整个网络, 能提高网络管理水平, 并为相关管理人员提供及时、有效、科学的辅助决策。

#### 参考文献:

- [1] 赵丽. 网络测试技术在校园网建设与维护中的应用[J]. 计算机技术与发展, 2007, 17(9): 174-177.
- [2] 曹庆华. 网络测试与故障诊断实验教程[M]. 北京: 清华大学出版社, 2006.
- [3] Han Jiawei, Kamber M. Data Mining Concepts and Techniques [M]. 北京: 高等教育出版社, 2001.
- [4] Ohanekwu T E, Ezeife C I. A Token Based Data Cleaning Technique for Data Warehouse Systems [J]. ACM SIGMOD Record, 2004, 8(2): 155-165.
- [5] Vassiliadis P, Sellis T. A Survey on Logical Models for OLAP Databases[J]. Communications of the ACM, 2002, 28(4): 164-169.
- [6] 张志军. 基于数据仓库的企业管理决策支持系统[J]. 计算机应用与软件, 2005, 22(6): 65-66.
- [7] 周大海, 安云哲, 夏秀峰, 等. 数据仓库系统中 Excel 文档数据的 ETL 方法研究与实现[J]. 小型微型计算机系统, 2007, 28(10): 1813-1816.
- [8] 张旭峰, 孙未未, 汪卫, 等. 增量 ETL 过程自动化产生方法的研究[J]. 计算机研究与发展, 2006, 43(6): 1097-1103.
- [9] Agosta L. 数据仓库技术指南[M]. 北京: 人民邮电出版社, 2000.
- [10] 杨学兵, 蔡庆生. 一种基于概念层次的分层规则挖掘算法研究[J]. 华中科技大学学报, 2001(9): 22-25.
- [11] 王大玲, 于戈, 鲍玉斌, 等. 基于概念层次的数据挖掘算法的研究与实现[J]. 计算机科学, 2001, 28(6): 90-93.
- [12] 黄玉明, 毛宇光. 数据仓库中多重粒度划分的层次编码解决方案[J]. 计算机技术与发展, 2008, 18(10): 1-4.

(上接第 238 页)

- [8] 杨钰. ITIL 中 IT 基础架构管理模型设计与实现[J]. 计算机技术与发展, 2007, 17(4): 250-253.
- [9] 丁云成. 基于 ITIL 的 IT 服务管理系统的应用[J]. 电脑知识与技术, 2008(7): 1266-1269.
- [10] 王小芳. 基于 ISO20000 的 IT 服务管理平台的研究与实现[J]. 计算机系统应用, 2009(5): 17-20.
- [11] 刘海峰, 连一峰. 基于 ITIL 的网络安全运营管理体系研究[J]. 计算机工程与应用, 2007, 43(9): 193-197.
- [12] 吴云峰. IT 服务级别管理在大型企业内部的实施[J]. 电脑知识与技术, 2009(15): 3790-3791.
- [13] 谢若承. ITIL 理念在 SAP 系统管理中的实践[J]. 计算机应用与软件, 2010(5): 203-205.
- [14] 周永伟. 基于 ITIL 的网络运维流程的改进研究[J]. 计算机与信息技术, 2007(10): 43-44.
- [15] Nelson R. Best Practices When Implementing ITIL[C]//MISA BC Fall 2004 Conference. [s.l.]: [s.n.], 2004.
- [16] Galup S, Dattero R, Qaun J J, et al. An Overview of IT Service Management[J]. Communications of the ACM, 2009, 52(5): 127-129.