

# 基于数据仓库的本地网网络管理决策支持系统

The Local Telephone Network Management Decision Support System Based on Data Warehouse

唐 红 张世璐

(重庆邮电学院 重庆400065)

**Abstract** Mining abundant information involved in network management has great significance on improving telecommunication enterprise management level and bringing into play investment benefit. This article presents the framework of a network management DSS based on network Management Information Base (MIB) with data warehouse technique. Network management DSS also adopts Internet information putting out techniques such as Web technique and PUSH technique, which transmit decision information easily and quickly to decision manager.

**Keywords** Network management, DSS, Data warehouse

## 1 问题的提出

电信网络管理的目标是最大限度地利用电信网络资源,提高网络运行的质量和效率,同时向用户提供良好的通信服务。要达到这些目标,就需要及时地分析网络运行情况、网络资源的配置等,而现有的本地网网管系统对这些分析处理的支持一直不能令人满意,这就需要管理信息库中的操作型数据和分析型数据相分离开来,并且将分析型数据按决策支持要求重组,建立起单独的分析处理存储环境。所以在电信网络管理中建立基于数据仓库的决策支持系统,对于达到上述目标无疑是有帮助的。

本地网网管系统包括本地电话网网管、长途电话网网管、传输网管以及“九七工程”等系统。在现行体制下管理所有处于本地网范围内的长途和本地交换网和传输网,以及相关的电信业务,实现网络的性能管理、故障管理、配置管理、计费管理、安全管理等功能。但是,由于种种原因,现有网管系统的功能较差,主要完成性能、故障等数据的收集和设备的集中维护功能,没能为网络管理决策者提供关于网络运行、经营、规划等必要的决策信息,没有充分发挥其应有的作用和效益。存在的主要问题有:

①功能比较简单,不能满足多层次网络管理的需要;

②收集到的大量数据仅用于网络的日常维护等事务性工作,未做深层次的挖掘,信息资源严重浪费,不能充分发挥投资效益;

③不同专业网管系统自成体系,具有独立的管理信息库,相互之间不能互通,数据不能共享,如网络运

行的数据仅用于日常维护,而运营数据仅用于对网络业务的支持,这些数据没有得到进一步发掘和综合,造成极大的资源浪费。

④未提供通用的用户查询界面,只能由网管中心的操作人员进行相关的查询、管理工作,尚无远程用户的接入,故当网络发生重大异常事件时,不能及时响应处理。

由于上述这些问题,我们提出了建立基于数据仓库技术、数据挖掘技术和联机分析处理的本地网网管决策支持系统来解决网管信息未能充分利用的问题。

## 2 系统结构

目前我国各本地网的网络管理平台已基本建成,定义了统一的数据模型,能够对本地网进行集中监控,实时收集网络运行数据。本地网网管管理信息库的数据主要有故障管理数据、性能管理数据、配置管理数据、资产管理数据、帐务管理数据和用户管理数据等。本文提出的网络管理决策支持系统以现有网管系统的管理信息库为主要数据源,建立关于网络运行、网络规划和网络经营的数据仓库,开发数据提取软件,选择适用的数据挖掘 DM、联机分析处理 OLAP 等工具软件,根据电信网络管理的特点建立决策模型,在本地网网管系统和局内办公自动化系统平台上,利用 Web、PUSH 等信息发布技术形成管理决策信息,使用浏览器作为用户界面,系统的体系结构如图1所示。

由图1可看出,该系统包括数据提取层、数据仓库层、分析处理层、信息发布层以及数据源、用户浏览器等六个层次,其实质就是数据仓库的建立、信息的挖掘和信息的发布三个部分。

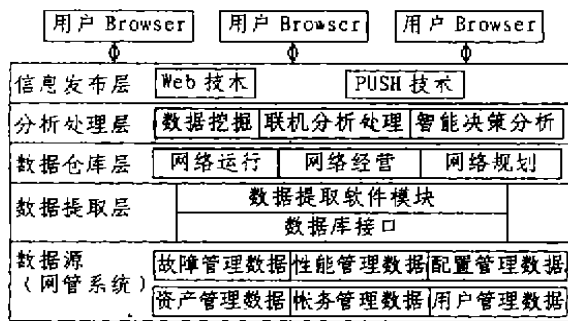


图1 网管决策支持系统体系结构

### 2.1 数据仓库的建立

数据仓库就是一个用以更好地支持企业或组织的决策分析处理的、面向主题的、集成的、不可更新的、随时间变化的数据集。数据仓库的数据是按主题组织的。本系统的主要数据源是网管系统的管理信息库(MIB), 主要包括故障管理数据、性能管理数据、配置管理数据、帐务管理数据和资产管理数据等。它含有大量的数据, 这就需要用数据仓库技术对这些数据进行重组, 从大量的维护数据中获取管理决策信息。并根据决策支持系统的需要, 以多种形式进行数据和信息的组织。本系统以呼叫详细记录、用户、话务等为主题来进行数据的组织。

这些数据源的数据是不同类型的, 所以要通过相应的方法和工具进行抽取和清洗以转换成相同类型的数据, 然后将这些数据进行分类、汇总。数据提取层就是开发数据提取软件来完成这项功能。

数据仓库为了使用户能方便地获取信息, 它还应包括查询工具、报表生成工具和多维分析处理(OLAP)。多维分析是指对以多维形式组织起来的数据采取切片、切块、旋转等各种分析动作, 以求剖析数据, 使最终用户能从多个角度、多侧面地观察数据库中的数据, 从而深入地了解包含在数据中的信息、内涵。

概括起来, 数据仓库在本地网网管系统中的应用主要有三种:

(1) 网络运行分析 网络运行分析的支持数据主要是性能管理数据和故障管理数据, 网络运行分析主要是分析网络话务流量、流向。对于突发性大话务量和某些话务拥塞引起的整体通信质量变差(如接通率下降), 或设备单元出现不太频繁或间断的差错等原因导致的网络运行质量变差, 而又不能被故障监测系统所发现时, 从网络运行分析上可以非常直观地看出网络实际运行的状态, 这样它就可以为网络调度提供决策依据, 从而及时发现故障或故障预兆, 分析故障发生原因, 减少故障和故障历时, 提高网络运行质量。

与网络运行分析相关的性能管理数据和故障管理

数据有: 安装电路数、可用电路数、试呼次数、接通次数、溢出次数、占用话务量、对端拥塞次数等如表1所示。

表1 网络运行分析的相关网管数据

字段名称	中文名称	字段类型
Ncct	安装电路数	int
Ncct	可用电路数	float
pegs	试呼次数	int
succ	接通次数	int
overflow	溢出次数	int
serl	占用话务量	float
dstcongest	对端拥塞次数	int

网络运行分析参数主要包括接通率、中继电路可用率、溢出百分比、中继电路群话务负荷、对端拥塞占用比。这些参数都可以由表1中的数据提取、转换得到。如接通率等于接通次数与试呼次数之比; 中继电路可用率等于可用电路数与安装电路数之比; 溢出比等于溢出次数与试呼次数之比等等。

(2) 网络经营分析 数据仓库系统可以根据用户号码、时间段和所在地点来分析全网话务分布、时间上话务量的变化趋势, 忙时的变化以及用户数据等, 比如地点、人口数量与话务量的关系, 时间与话务量的关系, 话务量的变化与现行资费政策的联系等, 从而发现潜在用户和业务量, 制定网络经营策略。

在分析上采用多维分析的方法, 数据进行多维数据库存储。如分析时间与话务量的关系时可以按日、年来进行分析。如表2、表3所示。

表2 一天中各时段话务量数据

		话务量			
		局向1	局向2	...	局向 n
时间	时段1				
	时段2				
	...				
	时段 m				

表3 一年中每月平均话务量数据

		话务量			
		局向1	局向2	...	局向 n
时间	1月				
	2月				
	...				
	12月				

例如如果通过这种多维分析直观看出时段1的话务量长期较多, 而时段2的话务量长期较少, 那么就可

以建议网络经营者制定一定的经营策略来减少时段1的话务量,增加时段2的话务量,从而平衡整个网络不同时段的话务量。

(3)网络规划分析 网络规划主要是起到优化网络资源,充分发挥投资效益的作用。它主要对性能管理数据、故障管理数据以及配置管理、资产管理方面的数据进行分析,按照理论与实际相结合的原则进行网络话务的预测、网络的规划、网络的调整和网络的扩容改造等。网络规划分析主要有以下应用:

- 话务异常分析:分析话务分布模型,检测超门限的拥塞状态,产生拥塞状态报表,表明拥塞程度和范围;

- 话务容量分析:显示网络范围的最忙时间和话务量,这提供了对网络资源的上界要求。

与网络规划分析相关的网管数据如下表所示。

字段名称	中文名称	字段类型
ncit	安装电路数	int
pegs	试呼次数	int
overflow	溢出次数	int
serl	占用话务量	float
time	设备购置时间	date
congest	交换机拥塞次数	int
seiz	占用次数	int

网络规划分析的参数主要有设备使用年限、电路利用率和交换机拥塞比。如电路利用率等于占用话务量与安装电路数之比;交换机拥塞比等于交换机拥塞次数与占用次数之比。

### 2.2 信息的挖掘

对于那些决策者明确了解的信息,可以用查询、OLAP 或其它工具直接获取,例如需要了解地理位置与接通率、话务量的关系和时间与接通率、话务量的关系,就要利用 OLAP 技术按时间维、地区维和话务量维来组织数据。而对于那些数据仓库中尚未被发现的知识,如隐藏在大量数据中的关系、趋势等信息就需要数据挖掘技术。以“告警信息”为例,网络的故障告警信息实时地由网管中心采集,并存储于数据库中,目前对于这些消息,管理维护人员从网管系统中获知某些地方出现异常,需要监测和处理故障问题。这些告警信息中,有些是严重的告警,必须马上处理;有些属于一般告警,可以暂缓处理;而有些告警信息则不需要处理,网络系统和设备可以自行恢复。这些信息都实时存储于数据库中,而实际上,网络运行中所发生的异常是有规律可以遵循的,某些告警会定期出现;某些告警类型之间具有一定的关联性。因此,运用数据挖掘技术,从告警数据库中可以发现网络运行的某些不为人知,却

具有一定特性的知识,以供管理维护人员进行分析决策。

### 2.3 信息的发布

在系统服务器端设置 Web 服务器,使用浏览器作为用户界面,采用 Browser/Server 体系结构。管理决策者可以及时、方便地查询到各类信息,如网络管理和维护人员随时可以查询当前网络运行状况,如话务分布情况,告警的类型、等级、发生的时间、地点等。对一些网络发生的重大异常和告警,利用 PUSH 技术,准确地将信息“推送”到相应的管理人员客户端。Web 技术和 PUSH 技术的应用将大大地提高网络管理效率和质量。

总之,从以上分析可看出本地网网管决策支持系统的特点主要有:

- (1)充分利用已建网管系统的资源,保证投资效益。本系统以已经建成的网管系统为基础,充分利用收集到的大量网管数据,并将它们合理而系统地组织起来,客观地为网管人员提供管理决策信息。

- (2)丰富的数据分析功能。本系统由于采用了数据仓库技术、数据挖掘技术和联机分析处理技术,可以多层次、多角度地分析和综合数据,挖掘出一些潜在的信息。

- (3)采用 Web、PUSH 等信息发布技术,使决策信息可以及时、方便、准确地传送到管理决策者手中。

## 3 应用系统

目前,全国三百多个本地网都已基本建立了本地网网管系统,很多电信局建立了局内办公自动化系统,网络配置、运行数据、用户数据等信息已经集中到网管系统的管理信息库中,这就为本地网网管决策支持系统的应用奠定了一定的基础。我们在已建的本地网网管与集中监控系统平台上建立了网管决策支持的初步应用系统,系统结构如图2所示。

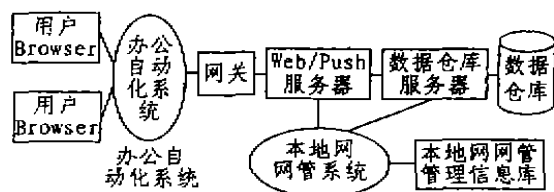


图2 本地网网管决策支持应用系统结构

该系统充分利用了已建成的本地网网管系统和办公自动化系统等网络资源和数据资源。它的数据源是本地网网管系统管理信息库中的网络配置、话务统计、故障告警、用户信息等数据。数据仓库服务器的采集单元定期从管理信息库中收集最新数据,并完成数据的清洗和重组任务。Web/Push 服务器响应用户的查询

请求,将信息发布到管理决策者手中,并利用 PUSH 技术将日常报表、重大异常或告警及时推送给相应的人员,用户通过局内办公自动化系统连接到系统的 Web/PUSH 服务器,并用网关将它们与网管系统的其它设备隔离,该系统的软件结构如图3所示。

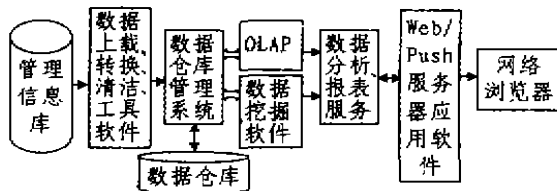


图3 本地网网管决策支持系统软件结构示意图

本应用系统中的软件体系是以数据仓库为核心,通过数据提取工具、数据转换工具、查询工具、报表工具、分析工具、数据挖掘工具等,满足用户对信息的各种需求。

在本系统中,我们主要考虑了对网络性能数据(话务量统计数据和网络配置数据)的利用。对管理信息库中的话务数据如忙时各中继群、目的码话务变化、每线话务量、电路可用率、电路溢出比、电路忙比等按日、月、年来进行分类、综合,并把这些信息以图表的形式

Push 给网络管理人员或供相关人员查询使用,网络管理人员从这些信息中可以分析网络性能、网络运行状况、网络负载变化趋势等信息。这些信息同时也为网络管理人员做出电路调度、阻塞控制、网络规划等提供决策依据。该系统已经用于实际工作中。

**结束语** 本文在分析了本地网网管系统现状的基础上,结合网管工作的实际需求,提出以本地网网管系统的管理信息库为数据源,建立基于数据仓库的网络管理决策支持系统的思路和实现方案,并初步实现了在本地网网管与监控系统基础上的应用系统并在实际工作中使用,取得初步成效。然而,研究工作仅仅是初步的,许多问题特别是相应的数据仓库模型和数据挖掘方法需要进一步探讨与研究。

## 参考文献

- [美] Hambergren T. 数据仓库技术. 中国水利水电出版社, 1998
- 王珊, 等. 数据仓库技术与联机分析处理. 科学出版社, 1999
- 王广清, 杨学良. 数据仓库技术及其在电信计费领域应用的探讨. 计算机工程与应用, 1999, 9
- 张伟民, 等. 数据仓库技术在电信管理网中的应用. 电子工程师, 2000

(上接第19页)

**定理8** 对于图4所示循环  $L$ , 如果能够并行执行, 则  $L \rightarrow \text{MIMD}$ 。

证明: 根据定义易证。□

**定理9** 对于任意循环  $L$ , 如果  $\rightarrow(L \rightarrow \text{SIMD}) \wedge \rightarrow(L \rightarrow \text{MIMD})$  成立, 则  $L \rightarrow \text{SISD}$ 。

证明: 根据定义显然。□

**结束语** 比较基于语句, 基于数组变量和基于数组元素的三种方法, 研究对象的粒度依次减小。基于语句的方法, 判断循环整体是否可并行执行比较容易。基于数组变量的解析方法可以研究多重循环的数据相关性。基于数组元素视图方法, 既有图直观的优点, 也具有代数理论形式化的支持; 既能研究循环整体并行化, 又能判定循环是否存在部分并行性, 还可讨论包含控制结构的循环程序的数据相关性问题, 不愧为是一种值得深入研究的方法。

数组元素时序层次图随着循环体语句数增多以及循环步数的增多, 结点数也将增多, 导致处理数据和运算量特别大, 给实际应用带来一定的困难。本文仅是对该方法进行了起步性研究, 还存在一些问题, 例如: 时序层次图标准形自动转换算法怎样构造, 如何在图上分割循环区间实现部分并行, 以及如何有效构造二重、

多重循环的时序层次图。这些问题是我们今后的研究工作。

## 参考文献

- Freund R F, Siegel H J. Heterogeneous processing. Computer, 1993, 26(6): 13~17
- Ekmečić L, Tarralja I. A survey of heterogeneous computing: concepts and systems. Proceeding of the IEEE, 1996, 84(8): 1127~1143
- Khokbar A A, et al. Heterogeneous computing: challenges and opportunities. Computer, 1993, 26(6): 18~27
- Zeng Guosun, Lu Xinda, Wang Jincun. Extracting Loop-Level Parallelism and Scheduling in Heterogeneous Computing Environment. In: The Third International Workshop on Advanced Parallel Processing Technology, Oct. 1999 228~230
- Zeng Guosun, Lu Xinda, Wang Jincun. Structure-based automatic extraction of the program heterogeneity. HPC-Asia 2000, May 2000, to appear
- 曾国荪, 陆鑫达. 异构计算中的负载共享. 软件学报, 2000, 11(4): 551~556
- Fan Zhibua. Discreteness of sequential hierarchy and its application. Scientia Sinica, 1988, 31(10): 1269~1279
- 曾国荪. 自动提取程序的异构特征实现异构计算. [博士学位论文]. 上海交通大学计算机科学与工程系, 2000