基于NETFLOW与SNMP的园区网流量监控系统

赵晓峰,徐义东
(安徽财经大学 网络中心,安徽 蚌埠 233041)

摘 要:随着园区网不断扩展,IP网络承载的数据流量越来越大,业务也越来越复杂。加上病毒泛滥、黑客攻击等因素,更加重了网络管理的负担。传统网络管理工具都是基于SNMP协议的,但SNMP只适合于简单的流量监控与设备管理,缺乏流量分析功能,当需要对流量进行分析或依据流量分析的结果实施网络管理时,传统工具就力不从心了。NETFLOW是一种可用于流量分析的协议,将其与SNMP协议结合使用,可有效弥补SNMP的不足,在充分发挥NETFLOW流量分析与SNMP流量监控与设备管理性能基础上建立的监控系统,可有效提高园区网管理的质量与效率。

关键词:NETFLOW;SNMP;流量监控;流量分析
中图分类号:TP393.18 **文献标识码**:A **文章编号**:1673-629X(2008)05-0168-04

Monitoring System on Campus Network Based on NETFLOW and SNMP

ZHAO Xiao-feng, XU Yi-dong
(Network Center, Anhui University of Finance & Economics, Bengbu 233041, China)

Abstract: With the rapid development of the campus network, IP network carries more and more pressure due to the growth of data traffic flow. The service is also more and more complex. Virus, hacker attack, make network management burden heavier. The traditional network management tool is based on the SNMP protocol. But SNMP only suits in the simple current capacity monitoring and the network devicemanagement, lacks the current capacity analysis function. When needs to carry on the analysis or the basis current capacity analysis result implementation network management to the current capacity, the traditional tool lacked the ability to do what one would like. NETFLOW is available in the current capacity analysis, it with SNMP protocol union use, can makes up SNMP effectively the insufficiency. Based on SNMP and NETFLOW establishment supervisory system, can effectively improve the quality and the efficiency which the campus network manages.

Key words: NETFLOW; SNMP; network traffic monitoring; network traffic analysis

0 引 言

随着园区网规模及应用不断扩展,网络管理的压力也日渐增加。对于网络管理人员来说,理解网络流量的内容、了解用户的网络行为是网络管理的重要内容,它为日常网络管理、未来网络升级与容量规划等提供了重要依据。网络管理中,NETFLOW协议是一种新兴的网络流量分析监控技术,是基于"流(FLOW)"的,利用流信息,可依据源、目地IP地址对数据包进行包个数、字节数的记录、统计。但单纯依赖NETFLOW,是无法定位IP地址物理位置的,如当想了解某一IP数据包来源于哪个接入交换机端口,单纯依赖NETFLOW无法实现,必须结合SNMP协议才能实现。

结合NETFLOW流量分析与SNMP流量监控功能设计的网络管理系统,可使网络管理性能达到最佳效果。网络管理人员通过这种管理系统,可以知道网络的通信状况、了解谁在使用网络、谁在什么时间访问了什么网站、何种类型应用和哪些计算机在消耗带宽,还可以用作故障诊断,确定影响网络运行质量的原因。另外,当网络发生安全事故(网络攻击、非法言论),安全部门要求协助调查时,提取的流记录历史信息可用来追踪用户上网痕迹,并可根据SNMP记录信息,查询要追查用户何时在何交换机何端口访问网络的。

收稿日期:2007-08-28
基金项目:安徽省自然科学基金(2006kJ017C)
作者简介:赵晓峰(1970-),男,安徽蚌埠人,研究方向为网络管理、网络安全。

1 相关技术概述

1.1 NETFLOW

NETFLOW是Cisco公司首创的为提高路由设备

路由转发能力而在交换、路由体系中采用的一种三层交换技术,目前已成为事实工业标准,被包括Juniper、Extreme、Foundry等大多数主流厂商路由器和三层交换机支持。其工作原理是:NETFLOW利用标准路由模式处理数据"流"的第一个IP数据报,生成NETFLOW缓存,随后同样的数据基于缓存信息在同一个数据"流"中进行传输,不再匹配相关的访问控制等策略,同时NETFLOW缓存包含了随后数据"流"的统计信息[1]。

"流"由以下7个关键域构成:(1)源IP地址(Source IP address);(2)目的IP地址(Destination IP address);(3)源端口号(Source Port Number);(4)目的端口号(Destination Port Number);(5)协议类型(Layer 3 protocol type);(6)服务类型(Tos);(7)数据流入的逻辑网络接口(Input Logical Interface)。每个活动的流在缓存中占有一项记录,当一个不同于现有记录特征的数据包进入时,就自动地为这一数据包在缓存中开辟新的流记录项。后续进入缓存的数据项,如果和已有的记录具有相同特征,其统计信息就会加到相应记录中去。缓存中的NETFLOW流记录信息,间隔一段时间会被刷新,如果路由或交换设备做了相应设置,缓存中的流信息可定期通过UDP包发往指定工作站。

流信息包括数据包三层以上信息,通过对收集的NETFLOW信息整理、分类、分析,可快速、方便地了解网络流量的内容。

1.2 SNMP

SNMP(Simple Network Manage Protocol)是1988年制定的,最初只想把它作为TCP/IP网络管理的临时解决办法。但此后,由于其简单实用而被业界广泛接受,因此SNMP成了事实上的计算机网络管理国际性标准[2]。

SNMP侧重于网络流量监控与设备管理,利用它,可根据设备端口来获取网络流量大小信息,如网络接口的输入/输出数据包数、字节数等,但具体每个数据包三层以上信息,通过SNMP没有办法了解。

因此利用NETFLOW对流量进行分析,再结合SNMP对具体流量的物理位置进行定位,并对设备进行管理,这样才能使两种协议的功能得到充分发挥。

2 监控系统设计

2.1 系统总体结构设计

本监控系统是一套基于WEB的流量监控、分析与管理系统,用于实现对网络流量的实时监控、网络流量物理来源定位、网络流量应用分析、网络设备管理等功能。

系统逻辑结构如图1所示,底层的NETFLOW采集模块收到包含NETFLOW数据的UDP数据包后,将处理后的NETFLOW信息导入MYSQL数据库。NETFLOW有多个版本,本系统使用的是流行的v5版。另一底层SNMP轮询、采集、管理模块用于对网络设备如交换机、路由器进行SNMP轮询与数据采集,并可在发现用户计算机感染蠕虫、进行黑客攻击时,关闭该用户所连交换机端口。WEB服务器端由4个模块组成,它们均用Java语言编写,以Servlet形式运行在Tomcat服务器上[3]。

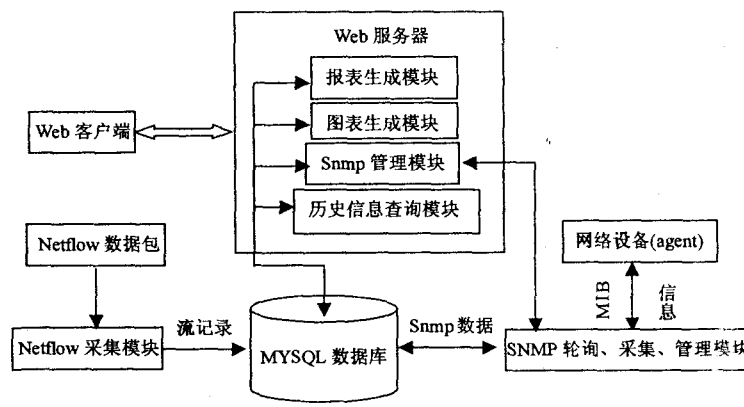

图1　系统逻辑结构图

所有模块即可全部运行在单台LINUX服务器上,也可把WEB服务分离,两台服务器一台运行Tomcat,该服务器操作系统可以使用LINUX也可以使用Windows,另一台LINUX服务器上运行NETFLOW采集模块与SNMP轮询、采集、管理模块及MYSQL数据库。

2.2 NETFLOW采集模块

一个完整的NETFLOW应用包括两部分组成:

(1)NETFLOW data exporter:采集流数据,并将统计结果输出,在一些高端的三层交换机(路由器)如cisco76、cisco65系列交换机上具有NETFLOW采集、输出功能,只要做相应设置就可将NETFLOW信息输出到指定的计算机。

(2)NETFLOW collector and analyzer:接收流数据,可对信息进行处理,并存入数据库。本系统NETFLOW采集模块使用Fullmer的Flow-tools。Flow-tools是LINUX下一套处理流信息的软件集合,flow-capture命令具有接收流信息功能,flow-print、flow-stat等命令用来查看、统计、接收flow信息,利用flow-export命令可将接收的流信息导入MYSQL数据库[4]。

2.3 SNMP 轮询、数据采集与管理模块

本模块基于 Adventnet SNMP API,通过 Java 应用程序完成对网络设备端口流量采集和存储。AdventNet SNMP API 是 Advent Network Management 公司推出的 Java 开发包,它遵守 Internet 的 RFC 规范,支持 SNMPv1,SNMPv2c 和 SNMPv3,它提供了一系列用于创建跨平台的 Java 和基于 WEB 的 SNMP 的 Applet 和 Applications 的 Java 类库(API),为构造网络管理产品和解决方案提供了良好基础,并简化了程序的开发过程。

本模块中,主要使用了 SnmpTarget 类及其方法:SnmpGet()、SnmpGetNext()与 SnmpSet()。SnmpGet()用于在轮询接入层交换机时,获取交换机接入端口 iflnoctets 与 ifoutoctets 值。通过 SnmpGetNext()对各接入交换机 MIBOID1.3.6.1.2.1.17.4.3 访问,可获取接入交换机自学习得到的接入计算机网卡 MAC 地址,通过 SnmpGetNext()对网关 MIBOID1.3.6.1.2.1.4.22.1.2 的访问,可实时获取整网 IP-MAC 映射信息。通过获取的这两种信息,就能唯一确定某 IP 对应 MAC 地址及该 IP 来源于哪个交换机端口,然后通过 JDBC 接口可定时将上面采集的 SNMP 信息存入 MYSQL 数据库。SnmpSet()用于对设置了读写团体名接入交换机的端口进行管理,当管理员通过 WEB 的 SNMP 管理模块发出关闭某一端口请求,Java 应用程序通过调用 SnmpSet()设置要关闭交换机端口 MIBOID1.3.6.1.2.1.2.2.1.7 值为 2,即可将其端口关闭,设该 MIBOID 值为 1,该端口将被打开。

2.4 数据库

在园区网中,同时在线的计算机有上千台甚至上万台,NETFLOW 流量数据非常大。高峰时 NETFLOW 流量记录甚至会超过每分钟几万条。因此向 MYSQL 导入 NETFLOW 数据时,应有选择地导入,例如只导入源 IP 地址、目地 IP 地址、源端口、目地端口、流中报文数、流总字节数这六项。从数据库中提取历史数据分析网络运行状况时,采用临时表等技术可以提高分析软件性能。数据库表结构的设计也同样需要进行优化处理,以达到节省存储空间和提高插入速度的目的。

因数据库中每天都会产生巨量数据,每天应定时对其清理,并且数据库中只保留近两个月的原始流记录与 SNMP 信息,而较早的记录将被聚合,只保留源 IP、目的 IP 等主要信息,以备进行月报与年报分析时使用。

2.5 报表生成模块与图表生成模块

报表与图表生成模块接收管理员查询并生成相应报表与图表。图表生成工具采用纯 Java 的 JfreeChart 来实现,可方便地与 Servlet 集成。JfreeChart 通过读取数据库的 NETFLOW 信息,可生成每个 IP 地址进出流量 TOP10 流量图、每种应用占整个流量百分比饼状图、月报、年报等各种图表。JfreeChart 通过读取数据库的 SNMP 信息,可生成每个交换机端口进出流量 TOP10 流量图。比照这些图表,可实时发现某一时刻,哪些 IP 地址处在带宽占用最高峰及这些 IP 地址对应于哪些交换机端口,还可查清这些 IP 地址流量大是哪些应用所造成。

2.6 SNMP 管理模块

SNMP 管理模块接收管理员发出的交换机端口关闭或开启请求。通过调用 SNMP 轮询、数据采集与管理模块中针对设备管理编写的 Servlet 程序,即可方便地实现交换机端口的关闭或开启。

2.7 历史信息查询模块

历史信息查询模块接收管理员发出的历史信息查询请求,通过对数据库中记录的 NETFLOW 数据及 SNMP 数据信息进行相关查询,可查询出任一时刻某一 IP 访问网络情况,及使用该 IP 的计算机网卡 MAC 地址和其连接交换机端口。但因数据库只保留近两个月的详细数据信息,使用该模块只能查询两个月内用户上网信息。

3 分布式监控系统设计

当园区网规模很大时,NETFLOW 流数据量将非常庞大,Cisco 的官方白皮书中指出 NETFLOW 本身流量可以占到网络总流量的 1.5%。加上对设备的 SNMP 数据轮询等功能,如果监控系统采用单服务器构架,服务器将不堪重负,这时必须采用分布式架构,才能保证监控系统正常运转[5]。

分布式构架系统逻辑结构如图 2 所示,图中假设园区网由两个区域构成,每个区域用一台支持 NETFLOW 功能的三层交换机作为核心交换机或汇聚交换机。NETFLOW、SNMP 数据采集模块及数据库分别运行在两台 LINUX 服务器上,每台 LINUX 监控服务器只负责从本区域三层交换机上采集 NETFLOW 数据,并只对本区域下辖的网络设备进行 SNMP 轮询、数据采集与管理。

分布式系统中,NETFLOW、SNMP 数据采集及数据库结构基本不用修改。主要对运行在 WEB 服务器上的各模块做修改,如在报表生成模块与图表生成模块中增加数据汇总功能,从区域 A 数据库与区域 B 数据库读取数据后,要先对取得的数据进行汇总,在汇总的基础上生成 TOP10 流量图、应用占总流量百分比饼

状图等。历史信息查询模块增加判断查询IP来源区域功能,当管理员输入要查询的IP历史信息时,先经过IP地址来源区域判断,确定IP来源区域后,再到相应区域的LINUX服务器数据库中取相应历史记录信息。

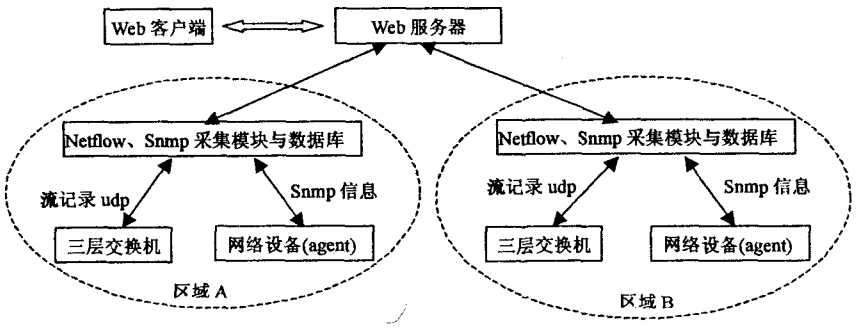

图2　分布式系统逻辑结构图

4　结束语

设计了一种基于NETFLOW与SNMP的网络监控系统。该监控系统要求不高,配置较为简单,只要园区网核心交换机支持NETFLOW功能,各接入交换机支持SNMP功能,即可实施,目前较大的园区网都满足这两项要求。

利用所设计的监控系统,能很好地对管辖范围内的网络数据流进行有效监控,给网管人员提供丰富的决策依据。通过实时监控,网管人员可以找到病毒、蠕虫和网络攻击等网络异常行为的源头,并采取有效措施进行阻断,以阻断异常流量的影响进一步扩大。本监控系统功能还不太完善,更多功能有待进一步研究与开发。

参考文献:

[1] Cisco NetFlow Services and Applications[EB/OL]. Cisco White Paper. 1999. http://www.cisco.com.

[2] Stallings W. SNMP网络管理[M]. 北京:中国电力出版社,2001.

[3] 熊齐邦,黄明哲. 基于NetFlow和异步服务的网络流量监测系统[J]. 计算机工程,2006(13):144-146.

[4] 何海涛,罗笑南,郭清顺. Netflow在边界网流量测量中的应用研究[J]. 计算机工程与应用,2004(11):11-14.

[5] Netflow Services Solutions Guide[EB/OL]. 2005. http://www.cisco.com/en/US/products/sw/netmgtsw/ps1964/products-implementation-design-guide09186a00800d6a11.html.

(上接第143页)

改进活动中总结以下经验:

(1)在应用CMM理论进行软件过程改进中,要根据软件企业的发展情况及其系统开发环境,从中找出制约软件企业开发效率和能力的关键因素,提出可行的标准和实践的步骤。根据软件企业的实力确定过程改进的各个阶段目标,不断总结、改进、提高,找出一条适合本软件企业的CMM实施路线。

(2)针对软件企业自身的特点,对CMM过程进行适当裁剪,对于CMM的模型与标准,不能生搬硬套,而应将其作为参考[5]。必须结合项目的情况和企业自身的特点、要求与现实条件,制订软件过程和选择实行改进的部分。在引进、消化、吸收的基础上需要自主创新,让CMM更实用化,形成自己的管理模式。

(3)充分认识改进过程本身就是一个规范的过程,需要循序渐进、逐步改进,因为软件过程成熟度的升级本身就是一个有生命周期的过程,而且全面引进CMM所涉及的范围非常广,要求人力、财力与设备资源的投入跟得上。在最初实施CMM时,软件企业可以试行某些关键过程域的一部分关键实践活动。并逐步规划出软件过程建立与改进的短、中、长期目标,分清轻重缓急,逐步取得经验。

(4)实施CMM对软件企业的发展有着重要的作用,CMM过程本身就是对软件企业发展历程的一个完整而准确的描述。软件企业通过实施CMM,可以更好地规范软件生产和管理流程,使企业组织规范化,提高软件企业的能力成熟度,改进软件的开发、维护过程,按时、按预算为用户提供高质量的软件,提高产品和企业的竞争力。

参考文献:

[1] 卡耐基梅隆大学软件工程研究所. 能力成熟度模型(CMM)软件过程改进指南[M]. 刘孟仁,等译. 北京:电子工业出版社,2001.

[2] Persse J R. CMM实施指南[M]. 王世锦,蔡愉祖译. 北京:机械工业出版社,2003.

[3] Caputo K. CMM实施与软件过程改进[M]. 于宏光,王家锋,等译. 北京:机械工业出版社,2003.

[4] 郑人杰. 基于软件能力成熟度模型(CMM)的软件过程改进[M]. 北京:清华大学出版社,2003.

[5] Raynus J. CMM软件过程改进指南[M]. 邱仲潘,等译. 北京:电子工业出版社,2002.