城域网流量采集系统解耦化研究及应用

Research and Application of Decoupling of Traffic Collection System in MAN

贾照娜[1]，朱友芬[2]，冯铭能[1]，王　欣[1]（1. 中讯邮电咨询设计院有限公司广东分公司，广东 广州 510627；2. 中国联通广东分公司，广东 广州 510627）

Jia Zhaona[1]，Zhu Youfen[2]，Feng Mingneng[1]，Wang Xin[1]（1. China Information Technology Designing & Consulting Institute Co.，Ltd. Guangdong Branch，Guangzhou 510627，China；2. China Unicom Guangdong Branch，Guangzhou 510627，China）

摘　要：

随着城域网流量的增长，流量采集专用系统的建设成本越来越高，某些运营商的设备能力已达不到考核要求。基于此，提出了针对城域网流量采集专用系统的软硬件解耦方案，采用X86通用服务器与纯软件相结合的方式采集城域网FLOW流量数据，并通过大数据平台实现各类应用分析；结合DNS日志，扩展实现了CDN流量流向分析。

关键词：

城域网；流量采集；解耦；大数据

doi：10.12045/j.issn.1007-3043.2019.09.018

中图分类号：TN915

文献标识码：A

文章编号：1007-3043（2019）09-0082-03

Abstract：

With the growth of MAN traffic，the construction cost of the dedicated traffic collection system is higher and higher，the equipment capability of some operators has failed to meet the assessment requirements. Based on this，it proposes a software and hardware decoupling scheme for the dedicated traffic collection system in MAN. By combining the X86 universal server with the pure software，the system can collect the FLOW traffic data of MAN，and it can analyze various kinds of applications through the big data platform. Combined with DNS log，the system extends its function of the CDN flow and direction analysis.

Keywords：

MAN；Traffic collection；Decoupling；Big data

引用格式：贾照娜，朱友芬，冯铭能，等. 城域网流量采集系统解耦化研究及应用[J]. 邮电设计技术，2019(9)：82-84.

1 概述

城域网流量采集系统是国内外各大运营商数据网流量采集的主要系统，它可实现全网流量采集、流量流向分析、异常流量检测和告警等功能。根据工信部要求，此系统必须建设，且随着城域网流量的增加，必须进行同步扩容。

目前全国大部分运营商的流量采集系统是由国外专用软硬一体化设备实现，设备供应商少、可选择范围小、价格昂贵且维护力量薄弱。随着宽带战略的推进，城域网流量年增长率达到30%，流量采集系统的建设成本越来越高，而运营商却无直接的经济收益。广东联通现有的流量采集系统由国外专用软硬一体化设备件实现，其设备能力已远远达不到考核要求。

基于此需求，本文提出一种软硬件解耦解决方案，采用X86通用服务器与纯软件方式结合，实现城域网FLOW流量数据采集，并通过大数据平台实现各类应用分析。在原有功能基础上，结合DNS日志，扩展实现CDN流量流向分析。该方案中的存储及分析平台利用大数据技术有效解决系统性能及功能的可扩展性问题，大大降低流量采集系统的建设成本。随着业务需求的增长，通过增加通用服务器即可实现性能

收稿日期：2019-06-29

的平滑扩充，通过增加软件模块即可实现功能的灵活扩展。

2 基于软硬件解耦的城域网流量采集系统

2.1 系统架构

软硬件解耦流量采集系统使用X86通用服务器及LINUX操作系统，采用纯软件方式实现对城域网FLOW流量数据的高性能采集，同时可实现对DNS、BGP、SNMP等其他数据源的数据采集。经过论证分析及测试，广东联通采用的系统总体架构如图1所示，数据采集完成后，通过HADOOP大数据平台进行存储及多维度分析，然后通过数据处理平台进行融合分析后，实现对应用功能的高度支撑。

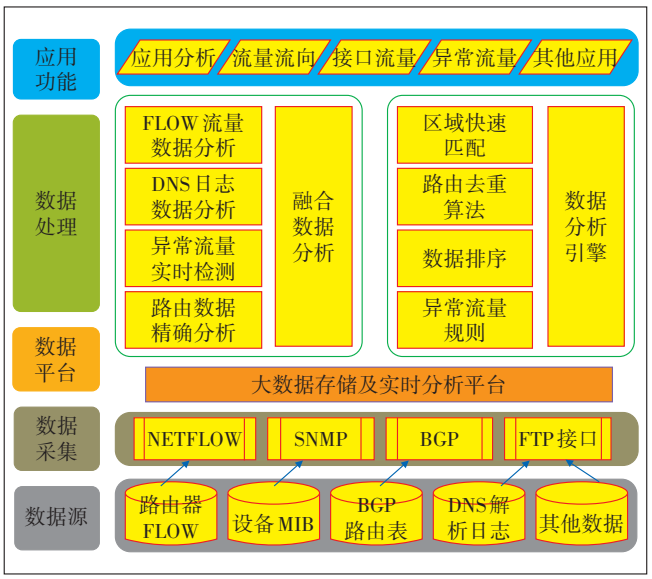

图1　系统总体架构图

2.2 部署方案

广东联通城域网流量采集系统部署方案如图2所示，分为流量采集服务器与数据分析服务器2部分，其中流量采集服务器主要负责采集数据，数据分析服务器主要负责分析数据并将结果呈现出来。

目前广东联通的数据采集接口包括FLOW流量数据采集接口、SNMP管理信息采集接口、BGP路由信息采集接口、DNS解析日志采集接口，也可根据业务需求开通与号线、IDC用户的数据采集接口。

3 流量采集实现技术手段

3.1 软硬件解耦

在X86通用服务器+LINUX操作系统的平台上部署流量采集软件，替代传统的软硬一体化设备，以实现城域网流量数据的高性能采集。

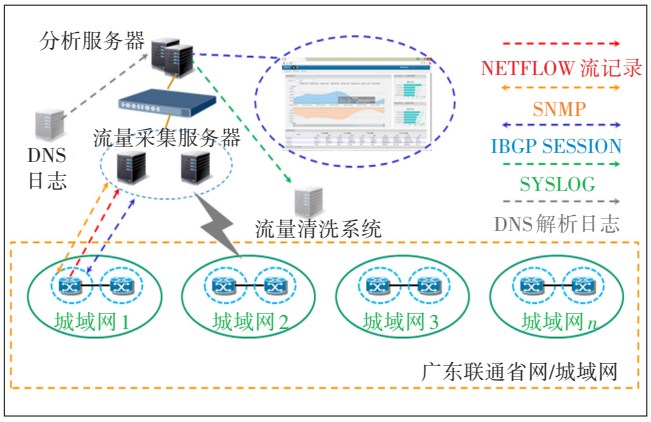

图2　广东联通流量采集系统部署方案

采集流程如图3所示，城域网设备在转发数据流的同时，根据设备的配置，实现对数据流的采样、统计，生成FLOW流量日志，并通过通信协议将FLOW流量日志发送至流量采集系统；根据城域网设备厂商及配置的不同，FLOW日志传送协议分为NETFLOW V5、NETFLOW V9、NETSTREAM、CFLOW/JFLOW；流量采集软件完成协议适配、FLOW日志字段解析并生成文本文件传送至大数据平台，由大数据平台完成后续的应用分析。

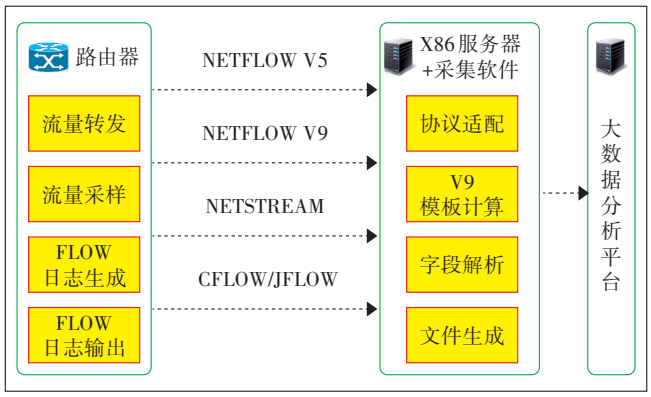

图3　城域网FLOW流量采集方案

3.2 大数据存储与分析

大数据存储与分析平台采用“HADOOP+SPARK+HBASE”框架，如图4所示，FLOW数据、DNS日志分别通过数据采集接口解析后保存在HDFS分布式文件系统，BGP路由数据使用QUAGGA并通过IBGP协议采集并保存至MySQL数据库。所有数据通过SPARK分布式运算平台进行分析后，保存结果至HBase数据库及MySQL数据库。

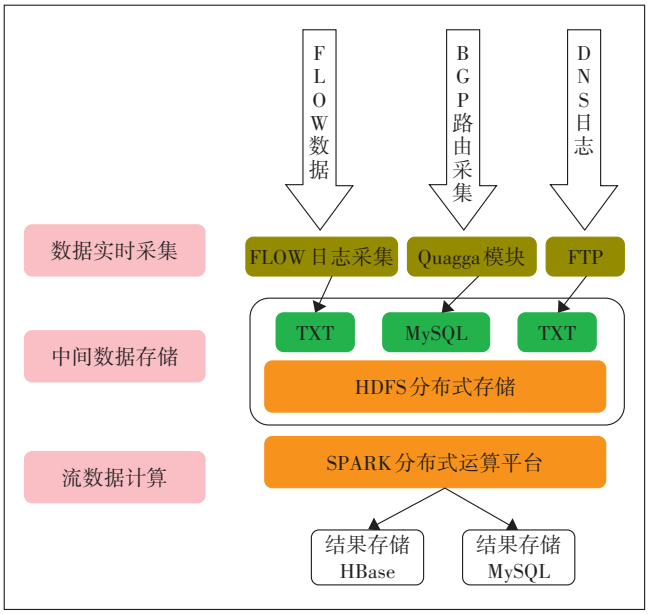

图4　大数据存储及分析平台

通过分布式数据存储和实时处理，系统性能和功能都具有高度的可扩展性。

4 应用场景

针对运营商网络中使用独立服务器资源的CP/SP/CDN服务商，结合DNS日志、FLOW流量数据、BGP路由数据，分析其特定区域的服务器流量流向分布或特定区域的客户至指定CP/SP/CDN流量流向分布，评估流量本地化效果，或者将分析结果作为CP/SP/CDN流量调度优化的依据。

在实际网络应用中，当某城域网用户持续稳定地访问某热点资源，而该热点资源又不在本城域网内时，则可以通过CDN的方式将该热点资源引入本城域网。考虑到业务需求和建设成本，初期可将CDN服务器设置在本城域网核心节点，并上联至CR，后续可根据业务发展情况、网络和机房资源，将CDN下沉至汇聚机房，并上联至BAS，如图5所示。这样一方面可以减少业务流量传输路径，减少丢包、时延，实现用户就近访问，保证服务质量，提升用户感知；另一方面可大幅减轻城域网和核心CDN节点的压力，有效支撑网络规划与优化。

5 结束语

广东联通率先使用X86通用服务器和纯软件相结合的城域网流量采集系统，该方案为行业内首创，打破了国外技术垄断，进一步提高了国家信息安全的可靠性，推动了国内企业在网络流量分析领域的发展。软硬件解耦后的城域网流量采集系统在性能、成本和扩展性方面明显优于传统专用系统，并可有效支撑面向应用的城域网流量多维度分析。

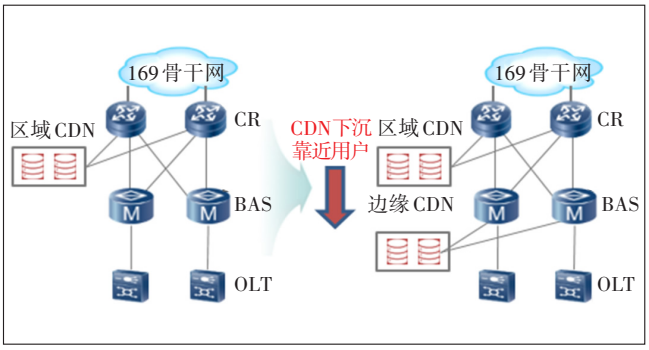

图5　CDN节点下沉示意图

参考文献：

[1] 谢喜秋，梁洁，彭巍，等. 网络流量采集工具的分析和比较[J]. 电信科学，2002，18(4)：62-66.
[2] 倪冰，陈运清. IP城域网流量监测方法及网络流量模型[J]. 吉林大学学报，2009，27(6).
[3] 陈宏. 城域网网络流量监测方法及实际应用场景[J]. 中国新通信，2014(24)：25-26.
[4] 曹平. 基于数据挖掘技术的网络流量分析预测系统研究[J]. 软件导刊，2012，11(5)：100-102.
[5] 汪文杰. 一种网络流量分析工具[J]. 数字技术与应用，2017(10)：77-78.
[6] 潘亮，张勇. 网络流量分析系统技术研究与实现[J]. 科技资讯，2008(1)：74-74.
[7] 卢佳佳，陈榕. IP城域网流量预警模型研究[J]. 计算机与现代化，2012(11)：194-196.
[8] 熊伟，孟新予. IP城域网流量预测模型分析[J]. 移动通信，2015(17)：63-67.
[9] 梁立京. IP城域网出口流量预测方法[J]. 电信技术，2014，1(10)：115-119.
[10] 程小芳. IP城域网流量流向分析[J]. 通信与信息技术，2012(5)：47-50.
[11] 林秀. NetFlow技术在IP城域网流量分析中的应用[C]// 中国通信学会信息通信网络技术委员会年会. 2003.

作者简介：
贾照娜，毕业于北京交通大学，工程师，硕士，主要从事数据网络的相关咨询设计工作；朱友芬，毕业于南京邮电大学，工程师，硕士，主要从事数据网工程项目管理工作；冯铭能，毕业于中山大学，高级工程师，博士，主要从事数据网络的相关咨询设计工作；王欣，毕业于华中科技大学，高级工程师，学士，主要从事数据网络的相关咨询设计工作。