东华流量分析系统
技术白皮书

东华软件股份公司

http://www.dhcc.com.cn

目 录

一、 前 言

随着信息技术和网络的深入发展，网络已经成为人们日常工作生活中不可或缺的信息承载工具。网络规模的日渐增长，网络中承载的业务也越来越丰富。企业需要及时的了解到网络中承载的业务，及时的掌握网络流量特征，以便使网络带宽配置最优化，及时解决网络性能问题。目前企业在管理网络当中普遍遭遇到了如下的问题：

- **网络的可视性**：网络利用率如何？什么样的程序在网络中运行？主要用户有哪些？网络中是否产生异常流量？有没有长期的趋势数据用作网络带宽规划？
- **应用的可视性**：当前网内有哪些应用？分别产生了多少流量？网络中应用使用的模式是什么？企业内部重要应用执行状况如何？
- **用户使用网络模式的可视性**：哪些用户产生的流量最多？哪些服务器接收的流量最多？哪些会话产生了流量？分别使用了哪些应用？

从这些企业管理网络中所经常遇到的问题来看，需要有一种解决方案能让网络管理人员及时了解到详细的网络使用情形，使网络管理人员及时洞察网络运行状况、及时了解网内应用的执行情况。

另一方面，伴随着互联网的正常应用流量，网络上形形色色的异常流量也随之而来，网络中经常出现的 DOS/DDOS 攻击，Red Code、SQL Slammer、冲击波、振荡波等网络蠕虫病毒泛滥对全世界网络造成的影响至今仍记忆犹新。各种网络扫描工具产生的大量异常网络流量大量占用网络设备系统资源（CPU、内存等）和网络带宽资源，使网络产生拥塞，造成网络丢包、时延增大，最终造成整个网络瘫痪。

为保障网络系统的正常运行，对于网络流量的统计、网络的安全控制就显得尤为重要。因此我们建议建设一套完整的流量管理平台，实现对流量系统的综合监控管理，并建设一个流量管理平台，规范信息中心的运行维护流程。

东华流量分析系统是东华软件股份公司结合网管领域的多年丰富经验推出的一套基于 xFlow（包括 NetFlow、NetStream、sFlow、cFlow、IPFIX 等）数据流的网络数据实时监控与分析产品。系统可以实现对大型网络的实时流量分析、历史流量统计、流量异常告警、流量趋势分析等功能，从而更好的发现网络异常流量、有效监控用户上网

行为，并为网络扩容投资提供科学的决策数据。东华流量分析系统采集对象通常是网络中的核心路由器和交换机，通过采集、监控网络第二至第四层的数据流量，实现 DOS/DDoS 攻击源和目标检测、蠕虫病毒监测、垃圾流量检测、异常流量告警、用户流量日志、网络带宽优化等。

二、 流量分析的目的

网络流量的复杂性给网络运营维护和业务决策支持带来了巨大挑战。一方面，业务决策人员需要详细了解网络带宽利用率情况和业务系统使用情况，提前为业务系统优化、升级扩容等方面提供解决方案，流量分析系统工具可以提供可靠的决策数据；另一方面，网络运维人员通常关心的问题包括两大类，一类是与网络和业务规划相关的问题，即关于网络流量成分组成以及地域分布的情况。另一类是与网络安全运营相关的问题，即关于异常流量的种类和数量、来源等情况。

流量分析的主要目的就是为业务决策人员和网络运维人员提供数据参考依据，为业务系统和网络运维提供全面的技术支持。下面分别详细列举每一类别所涉及的具体问题。

2.1业务决策支持

与业务决策相关的问题基本是围绕几个方面的内容：全网掌握网络带宽利用率、

业务系统流量的来源与去向、流量的组成成分、流量的变化趋势。例如：

- 从子网A 到子网B 的业务流量数据是多少？
- 各业务系统流量分布状况如何？
- 各个子网间的流量是否平衡？
- 网络出口业务系统的流入、流出流量是多少？
- 上联链路的负载是否均衡？
- 过去几个月的流量变化趋势如何？网络带宽是否足够？预计什么时候需要扩容？
- 内部网络到外部各个分支的流量的比率?
- 如何提升IT服务质量和能力？

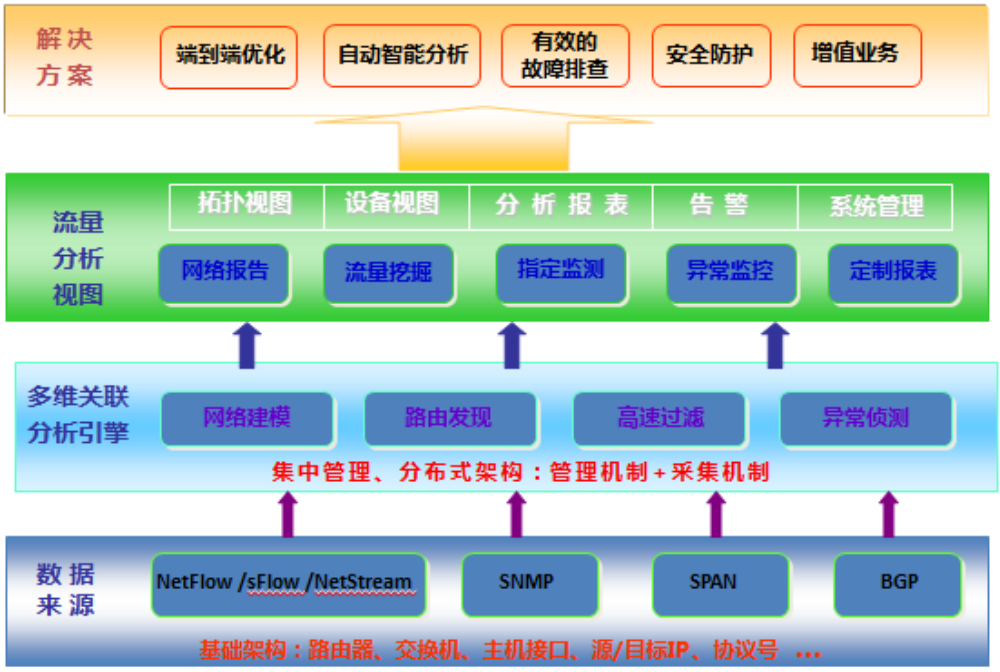
2.2网络运维支持

对网络的安全运营最大的威胁为异常流量和攻击。因此与安全运营相关的问题大都是围绕故障分析和异常流量展开的。运维人员最关心的问题有：

- 谁在访问我们的网络，是否属于攻击？
- 异常流量有多大？都是什么类型的攻击？
- 攻击流量的来源是哪里？
- 网络内部哪些流量被攻击？
- 网络内部是否有向外的攻击流量？
- 异常流量是由于什么引起的？
- 出具当前网络运行情况报告？

三、 流量分析的原理

任何IP网络流量分析的过程都可以大概分为三个环节：数据采集、数据分析、结果呈现。下面分别从这三个环节讲述流量分析的原理。



3.1 数据采集的机制

数据采集的大体上可以分为NetFlow/NetStream/sFlow、SPAN(镜像)、SNMP三种方式，这些方式是与设备相关的。即某些设备只能支持某一种或几种采集方式。每种采集方式有其固有的优势和局限。

东华流量分析系统采用了NetFlow/NetStream/sFlow和SNMP、SPAN、BGP相结合的流量数据信息采集方式，做到全面准确精细的数据捕获，从而为分析、统计、告警及报表等提供了完备的原始数据。

■ NetFlow 与 SNMP 技术的对比

	SNMP 轮询	NETFLOW 采集
采集端口流量统计	是	否
采集端到端流量	否	是
采集业务层流量	否	是
采集完整用户业务流量	否	是
消耗网络设备资源	较小	较多（但对高端设备性能影响不大）
适用电路	各种速率	各种速率
适用范围	网络各个层次	网络汇聚层和核心层
采集数据全面程度	较少	较全面

3.2 流量数据所含的信息

无论是那种类型的流量数据，包含的信息都可以分为三类：空间信息、时间信息、技术指标信息。

- 空间信息是流量发生的地点，包括：路由器、物理端口、IP地址（段）、AS号、地域名称等。
- 时间信息是流量发生的时间：用分钟、时间片、小时、日、周、月、年 来度量 。
- 技术指标提供流量的业务特征的信息：应用类型，TCP-flag, ToS, 包大小……

这样流量数据实际上就构成了一个多维数据集。采集的原始流量记录包含以下内容： IP、协议+端口号、接口、相邻自治域、原始自治域、TOS值、 Next Hop、TCP标记等。

- 流出或流入指定设备端口的数据；
- 源或目的IP地址指定的数据；
- IP 包内高层协议指定的数据；
- 指定高层协议端口的数据；
- 指定数据包大小的数据。

3.3 流量统计分析的过程

所谓流量的统计分析过程，就是在指定的时空范围内，统计流量对不同维度的分布。例如，查看某一天之内，某个路由器上各个端口上的流入、流出的总流量，就是查看指定时间 范围内流量对路由器端口这个空间维度的分布。再例如，查看某段时间内某个端口上流量中各种应用所占比例就是查看指定的时间范围、指定地点流量对应用这个技术指标维度的分布。

流量分析过程还可以用另一种方式来解释。前一节中，我们把流量数据看作了一个多维数据集，流量分析的过程就是根据实际需要从不同的角度去透视这个多维数据集。

3.4 异常流量检测的原理

异常流量的检测分为三个步骤：检测指标实测值的计算，检测指标基线值的计算，

实测值与基线值的比较。

每一种检测指标都对应一种或可能的几种攻击。也就是说，有的检测指标是专门检测某一种特定的异常流量的。而有的检测指标出现异常时，则只能判断存在几种可能的异常流量。这种指标就是非特异性指标。

每种检测指标都有自己的基线，但基线的算法是类似的。通常基线算法有两种，一种是周期性基线，一种是移动窗口基线。如果检测指标的正常值的变化趋势有明显的周期性，则建议采用周期性基线。如果检测指标的正常值没有明显的周期性变化，而且在一个较小的范围内波动，则使用移动窗口基线效果比较好。

系统支持丰富的异常检测方法，包括：

- **DDoS攻击**
  - SYN Flood
  - UDP Flood
  - ICMP Flood
  - ACK Flood
  - DNS Query Flood
  - Http Get Flood
  - LAND Flood
  - IGMP Flood
  - TCP Flag NULL
  - TCP Flag误用
  - Protocol NULL
- **流量超常**
  - bps超常
  - pps超常
  - 会话数超常
- **蠕虫事件**
  - Code Red
  - 硬盘杀手
  - SQL Slammer
  - 冲击波
  - 冲击波杀手
  - 震荡波
  - 邮件蠕虫
  - WinNuke攻击
- **流量分布异常**
  - 源地址分散度异常
  - 目的地址分散度异常
  - 端口分散度异常

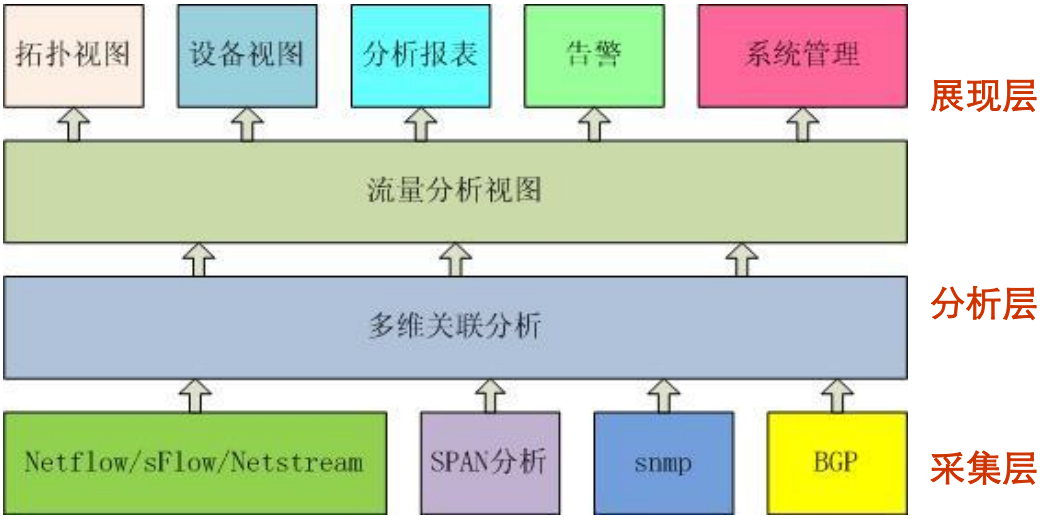
四、 流量分析的技术实现

目前网络流量分析产品的种类非常多，但是系统架构基本上是一致的。流量分析系统属于网络管理软件，因此在系统架构上仍然遵循网络管理软件系统架构的一般原则。即整个体系分为三层：采集层、分析层、呈现层。

采集层——采集层直接从网元设备获取原始流量数据并进行解码，把解码后的数据交给分析层进行分析计算，通常还会保留一份解码后的流量数据记录的硬盘。

分析层——分析层从采集层获取解码后的流量数据记录开始进行分析计算。通常在计算前需要进行数据的归一化处理。

展现层——展现层通过数据接口获取统计分析的结果并呈现在用户界面上。



东华流量分析系统采集对象通常是网络中的核心路由器和交换机，通过采集、监控网络第二至第四层的数据流量，实现 DOS/DDoS 攻击源和目标检测、蠕虫病毒监测、垃圾邮件检测、带宽成本分析、流量计费、用户流量日志、网络带宽优化等。

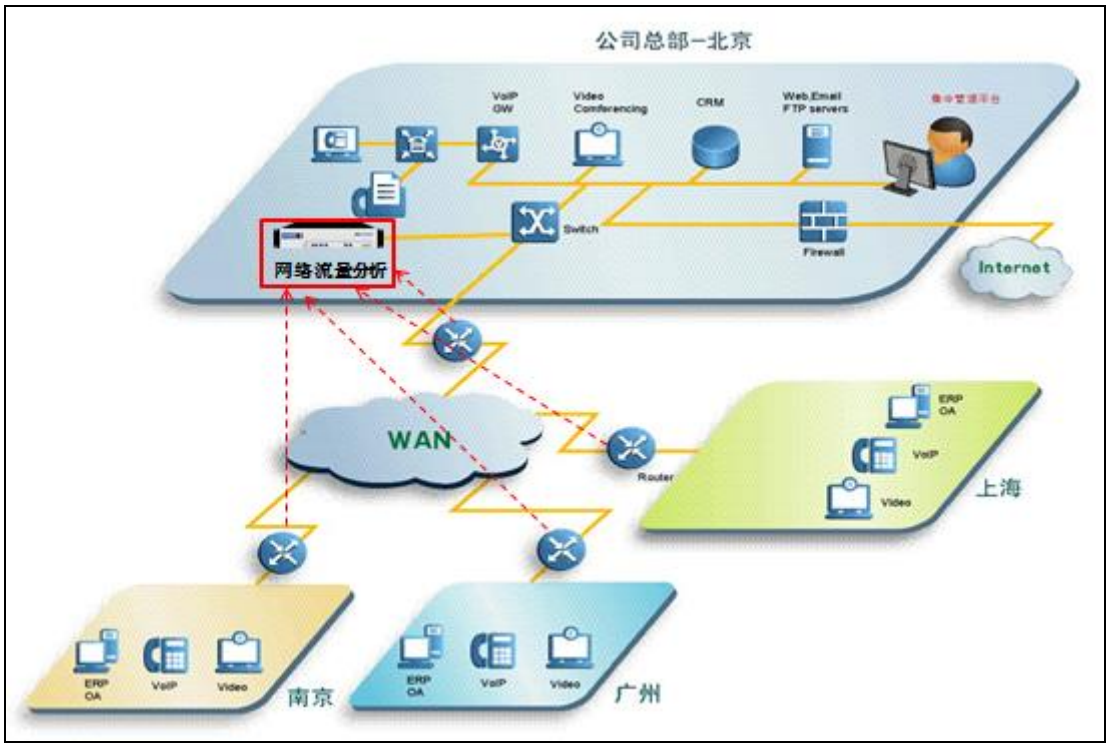
东华流量分析系统实时收集来自多个网络设备的符合 NetFlow V1/5/7/9、NetStream V5/9、sFlow V2/4/5、cFlow、IPFIX 等标准的详细流量记录，并对采集数据作复杂的运算与分析，根据预先设定的条件过滤 NetFlow 输出的流量数据，将符合监控条件的流量做分类与统计，并将结果存入数据库中。使用者可以设定监控条件，通过 Web 界面去查看网络流量的详细信息，详细流量数据经过整理后根据预定义的条件生成 TOP N 报告，此外 NetFlow 的原始数据(Raw Data)可以永久储存下来，以方便做历史流量分析。

五、 东华流量分析系统

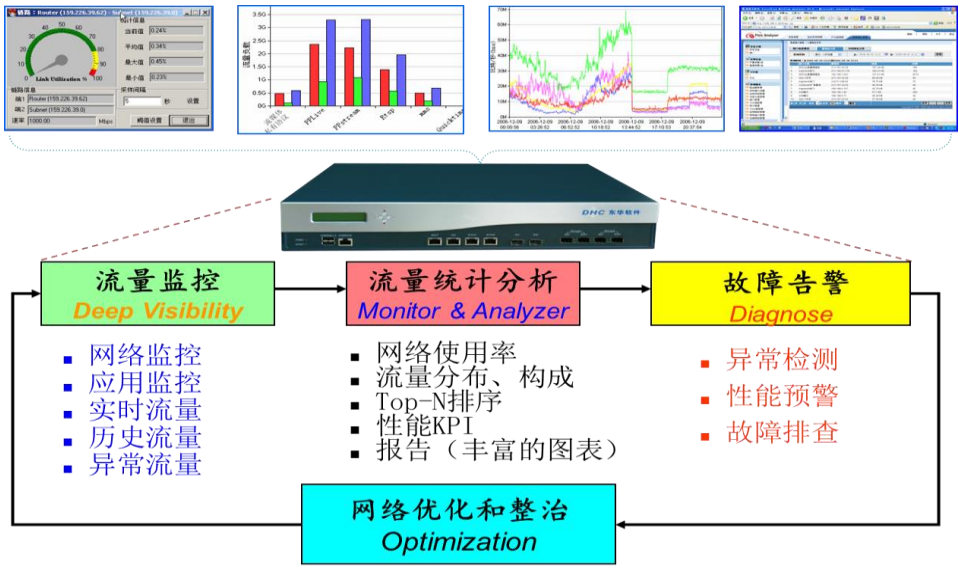
5.1 系统介绍

东华流量分析系统是东华软件股份公司结合网管领域的多年丰富经验推出的一套基于 xFlow 数据流和 SNMP 的网络数据实时监控与分析产品。系统可以实现对大型网络的实时流量分析、历史流量统计、流量异常告警、流量趋势分析等，从而更好的发现网络异常流量、有效监控用户上网行为，并为网络扩容投资提供科学的决策数据。

东华流量分析系统采集对象通常是网络中的核心路由器和交换机，通过采集、监控网络第二至第四层的数据流量，实现DOS/DDoS攻击源和目标检测、蠕虫病毒监测、异常流量清洗、用户流量日志、网络带宽优化等。



东华流量分析系统功能包括网络流量监控、流量超量预警、趋势分析、排序分析、实时分布、报表查询等。系统采用 Web 的界面，管理人员可以设定监控条件，通过 Web 界面去查看网络流量的详细信息，详细流量数据经过整理后根据预定义的条件生成 TOP N 报告，此外 xFlow 的原始数据(Raw Data)可以永久储存下来，以方便做历史流量分析。



- 采用独立的硬件结构平台，不需要增加其他的软件；快速安装、配置，可以独自完成流量的采集、过滤、分析和数据的存储；
- 高性能、可扩展的分布式体系结构；
- 高可用性设计；单台设备处理能力高达 160000flow/s；平均无故障运行时间(MTBF)大于 60,000 小时；实时流量数据刷新周期 30 秒，历时流量信息统计和报表生成时间 1/2/5 分钟；
- 支持系统弹性扩展和升级，满足用户网络增长需求，保护用户原有投资；
- 部署简单、灵活，通过旁路的方式接入网络，不会影响用户现有的网络结构和系统应用。可集中部署，或分层分级部署。

5.2 功能及其特点

系统主要功能：

- **流量实时采集监控**
- **流量报表展现**
- **全网流量关联性分析**
- **路由负载分析**
- **基线分析及告警**
- **异常流量监测及缓解**
- **拓扑视图**
- **热点网站的分析**
- **网内热点的流量分析**
- **内容源流量监测**
- **流量建模**
- **提供大客户增值业务服务**

功能特点：

系统架构

- ●采用独立的硬件结构，不需要增加其他的软件；
- ●快速安装、配置，可以独自完成流量的采集、过滤、分析和数据的存储；
- ●支持高可用性设计，可以同时部署两台设备接收流量数据；
- ●支持系统弹性扩展和升级，满足用户网络增长需求，保护用户原有投资；
- ●部署简单、灵活，通过旁路的方式接入网络，不会影响用户现有的网络结构和系统应用；

操作管理

- ●采用中文 WEB 操作界面，面向普通网络管理者，实现简单网络管理；
- ●强大的在线文档帮助功能，不需要系统培训就可以配置和使用；
- ●支持多用户分级权限管理，可以支持同时多人在线使用；
- ●支持设备分组管理，可以为不同的管理用户分配特定的管理设备和管理权限或为特定的管理设备分配多个管理人员；

流量监控

- ● 支持 NetFlow V1/5/7/9、sFlow V2/4/5、Netsream V5/9 等标准网络流量信息协议；
- ● 分布式的部署方式，能够实时采集和分析全网的流量数据；
- ● 通过接收网络流量数据，同时结合 SNMP 协议对网络设备运行提供全面监控、分析；
- ● 可以同时接受多种网络设备的流量数据，并支持实时转发流量数据，使用者根据要求快速扩展并进行综合性的统计分析；
- ● 支持镜像方式采集数据，进行网络流量收集并对网络中的流量进行分析。
- ● 系统缺省按照路由器的接口进行流量数据的监控和分析；
- ● 可以实时监控路由器和交换机中各个端口的通断情况并报警；
- ● 用户可以自定义特定子网范围，进行临时及长期的精确流量监控；
- ● 内置硬盘和数据库系统，支持大容量数据存储，用户可随时进行网络监测和分析，可以设置原始数据的保留时间，过期资料可以根据策略自动删除；

流量分析

- ●提供完善、丰富的报表展现，报表格式包括折线图、柱状图、圆饼图等，同时进

行 HTML、excel 等报表输出。

- 具备流量排序功能，可做流量累积统计或实时流量分析，流量排序支持自动设定，可按时自动生成 TOP N 排序报表；
- 可以对网络中的流量根据源地址、目的地址、源端口、目的端口条件进行详细查询，帮助用户对网络中的流量进行计费和网络监控功能。
- 内置强大的实时流量分析功能对于网络攻击的瞬间可做精确的来源掌握。
- 具备灵活的扩展能力，支持外部磁盘或 NAS 存储设备，支持大型网络环境中的实时、海量流量分析；

异常流量告警

- 具备异常流量告警功能，使用不同的颜色表示不同的告警级别，可以选择使用 EMAIL 或与网管系统结合的其他方式通知网管人员；
- 实现与 ITIL 事件、问题流程紧密集成，产生服务级别等级（SLA）报告、可用性及容量报告；

5.3 流量实时采集监控

系统实时收集网络设备xFlow流量信息，支持NetFlow V1/5/7/9、sFlow V2/4/5、Netsream V5/9等标准网络流量信息协议，对于不支持NETFLOW协议的网络设备可以采用镜像方式采集数据。监测互联网络出口、核心网络、特定子网间、总部与分支机构广域网、关键服务器和主机、端到端应用等流量信息及路由器和交换机中各个端口的通断情况等，为网络管理人员提供最准确的性能数据，诊断网络中的异常。采集的原始流量记录包含以下内容：IP、协议+端口号、接口、相邻自治域、原始自治域、TOS值、 Next Hop、TCP标记等。

- 流出或流入指定设备端口的数据；
- 源或目的IP地址指定的数据；
- IP 包内高层协议指定的数据；
- 指定高层协议端口的数据；
- 指定数据包大小的数据；

系统支持特定端口应用的原始流量数据的保存，用于历史的查询和痕迹保留。例如：用户可以选择保留互联网出口所有基于 80 端口应用的流量原始信息，保留时间根据用户的需求和设备容量进行定义。

5.4 流量分析

系统采用先进的比对和智能数据分析技术对网络总体流量、input/output 流量、应用类型、源地址、目标地址、详细信息进行流量过滤以及逻辑运算。它集成了多项流量分析功能，将短期实时监控、中期流量分析、长期流量统计功能紧密集成于单一系统，从而快速发现网络中的病毒扩散、大量的数据下载、IP 地址扫描行为、服务器某服务受到密集访问等网络异常流量问题，快速定位和解决故障。

用户根据实际网络状况分析互联网出口、核心局域网络、广域网、子网间、核心应用服务器的流量信息，自动过滤、分类出入流量，并产生各项深入分析报表：如流入、流出及进出本域网络的总量、应用协议、传输协议、端口使用、应用分布情况、自治域间的流量等等。

实时分析可以根据网络总体流量、input/output 流量、应用类型、源地址、目标地址、详细信息产生各种图表和 TOP N 报告。

流量实时查询可以快速反映网络中瞬间的流量状态，定位瞬间异常流量的产生及影响范围，通过TOP N排名报表，了解网络内实时数据的排名情况、协议类型、应用分布等，及时发现并准确定位异常流量的来源和目的及其流量的细节特性，网管人员可以及时排除网络故障，有效防止网络攻击。

通过对网络中一些特定流量的长期监控，有助于网管人员了解网络的流量模型和趋势，逐步形成网络流量的基线数据，同时结合可用性和应用性能为用户提供准确的服务

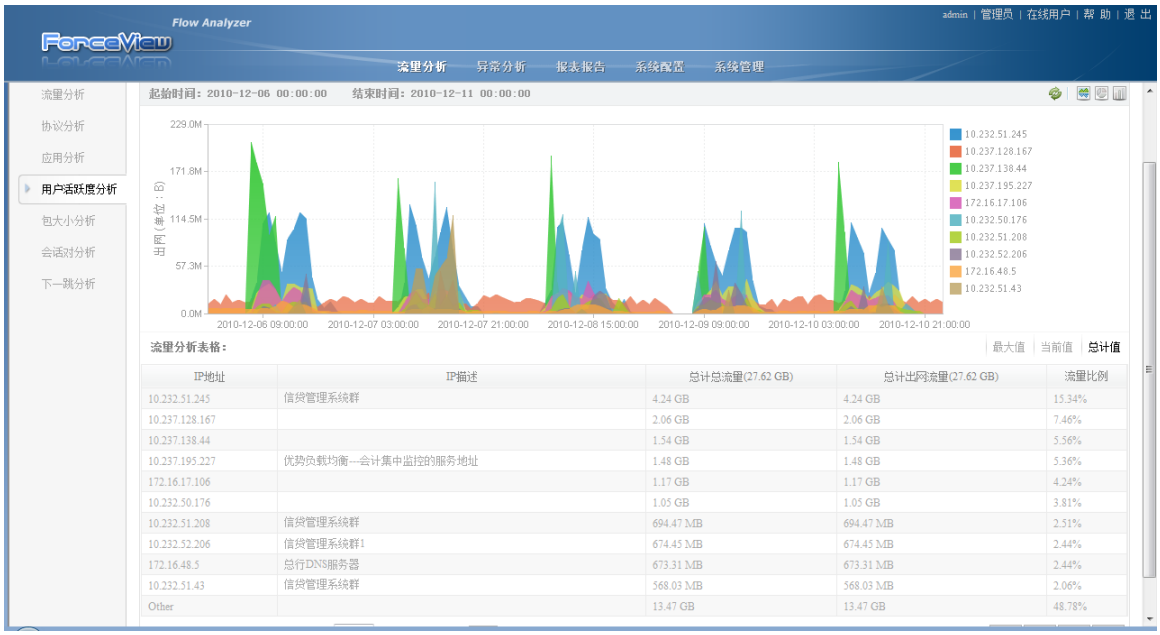
级别协议（SLA），为网络的优化以及业务的发展提供流量数据来源。对重要用户、业务、或可能病毒流量的长期实时的监测，可以准确把握并预防网络带宽拥塞的形成，通过快速定制ACL及其他安全措施，防止网络出现大范围阻塞，提供网络整体可靠运行。

5.5 流量报表展现

在数据的图形输出上，系统支持按时间查看数据流，可以生成日报表、周报表、月报表、年报表或任意时间报表，并使用TopN排序方式展示各数据分量的即时比特流、包流量、Sessions对话数；

网管人员可以建立自己的网络流量视图，用于监控互联网出口、核心网络、关键服务器、广域网络、VLAN之间的流量，可以按照以下字段分类显示原始数据及汇总数据：IP、协议+端口号、接口、相邻AS域、原始AS域、TOS值、 Next Hop、TCP标记等。

在数据的表格输出方式上支持将某时刻原始数据进行全面展示，允许在原始数据基础上进行排序、归类、过滤等分析操作；用户可以按照管理设备的实际连接单位进行自行定义，实现定制的客户化管理。



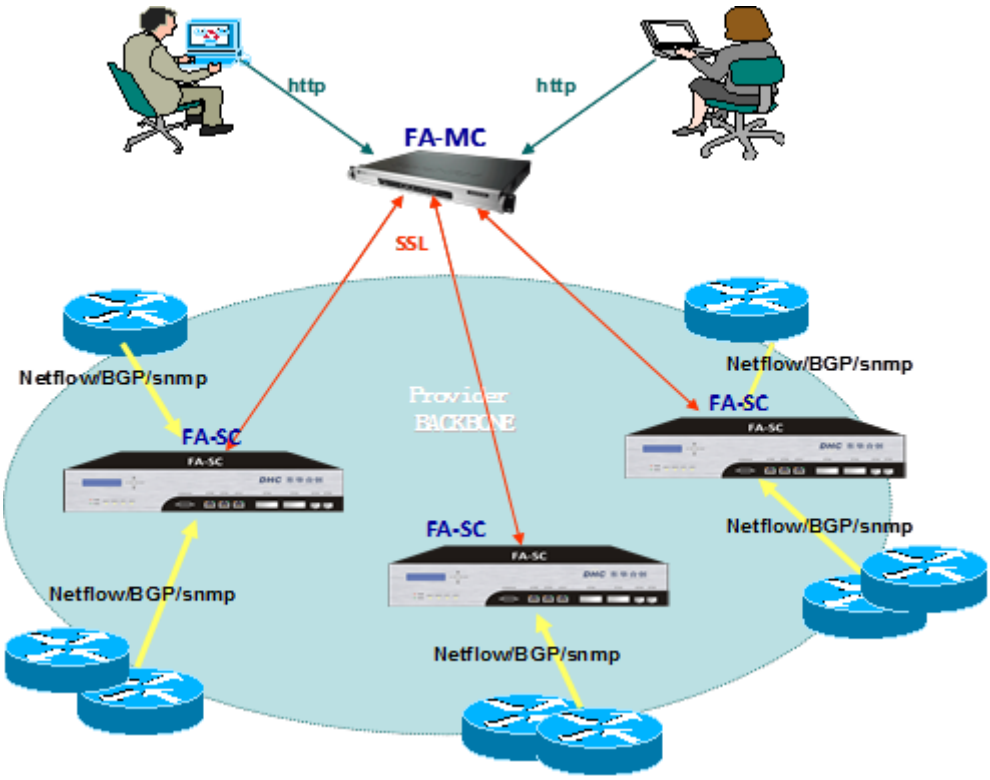
5.6 异常流量告警

提供灵活的告警管理功能，用户可以自行定义告警模板，根据已知的攻击和病毒端口，定义告警列表，支持邮件告警，同时结合网管系统的告警功能进行其他方式的告警。

根据长期监控的流量数据建立流量基线，并为基线流量设置最高超出的阀值及持续时间，一旦网络中出现超过流量基线的异常流量，就立即进行告警，并根据预先定义的告警规则进行预处理，生成异常流量相关报表，同时追踪异常流量的来源及类型，配合流量实时查询功能，获取异常流量特征的网络某一特定时间的实时流量数据，快速定位异常流量问题。

5.7 系统部署

■ 分布式部署，集中管理



5.8 系统优势

- 全面的流量采集机制，支持 NetFlow、NetStream、sFlow、cFlow、IPFIX 等协议，兼容 CISCO、华为、JUNIPER 等设备；支持 SNMP、BGP、SPAN 的流量信息收集。
- 专业化的独立硬件平台，电信级的可靠性与可用性，强大的处理性能（高达 16 万流/秒的处理能力）
- 非侵扰式的部署方式，即连即用，无需增加其他硬件探针或软件。部署简单灵活，可快速安装配置。
- 可扩展的分布式体系结构（流量采集机+分析管理机），支持系统弹性扩展和升级，满足用户网络增长的需求。
- 采用中文 WEB 界面，最适合国内普通网络管理者，实现简化网络管理；支持多用户分级权限同时使用。

- 灵活多样的监测分析引擎，智能的自动关联分析与异常检测机制；多种形式、多维度的报表展现。
- 能够提供定制化的 API 及二次开发，以满足客户不同的需求。
- 与流量管理系统 FS 组成强大的流量精细化解决方案。
- 与异常流量清洗系统 FG 组成强大的异常流量监测及清洗解决方案。

5.9 系统效益

通过网络流量监控分析技术，采用先进的系统工具做分析，由宏观面自上而下，掌握网络流量的大方向、由微观面自下而上，追查流量异常的关键因子，藉由主动/自动执行流量异常的抑制，使网络运作顺畅，智能地控制网络流量，并以支持客户未来提供丰富的差异化服务策略和保证机制为目标，通过这种梳理和协调来激发更多更好的网络应用。通过面向网络、面向用户和面向核心业务的网络流量分析和监控系统，全面掌握网络带宽资源、关键业务应用流量、用户访问行为等信息，优化网络及业务部署，从而改善整体网络的管理水平，提高网络传输效率和网络安全性，降低总体网络的运行维护成本、提升用户的网络业务服务能力。

通过对网络流量的监测管理，对系统进行全面业务应用情况的监测，保障网络稳定运行及关键业务的 QoS，这种系统管理办法，可以有效帮助科技部 IT 管理人员从“救火队”模式中摆脱出来，从而会节省大量的人力资源，并大大的提高工作效率。对网络系统的优化，实现网络从“无序”到“有序”，数据流清晰可视。这不仅可以营造一个健康高效的网络系统平台，而且对于提升用户的管理水平，提高网络系统利用率，增加用户的满意度，保障客户资源快速增长，实现利润最大化等方面有着重要而深远的意义。

同时，东华流量分析系统已具备农业银行、民生银行、华夏银行、安徽农信、江西农信、昆仑银行等多个金融用户案例，比较适用于金融行业用户。同时，在电力行业、电信运营商、政府等多个行业使用，流量分析系统已成为大中型网络流量分析的主要工具。

六、 结 论

东华流量分析系统是东华软件自主研发、拥有多项技术专利的产品，其技术处于国内外领先水平。此项目组拥有十几名博士，多名研究员、副研究员，东华软件新一代互联网网络监控分析系统研发及产业化项目被列入国家发改委 2008 年新一代宽带及网络通信产业化专项。

东华流量分析系统将帮助网络运维管理人员解决以下问题：

- 网络利用率如何？什么样的程序在网络中运行？主要用户有哪些？
- 哪些客户、流量占用大量带宽？如何提供端到端的流量监控？
- 如何预警将要发生的网络拥塞？
- 当网络中出现异常流量引发网络故障时，如何快速定位问题源？
- 如何验证 Qos 策略是否生效？
- 如何为优化网络结构、降低网络运行成本提供依据？
- 如何为网络投资升级提供准确的数据依据？

东华软件股份公司于 2003 年推出自主研发的网络流量分析与网络流量控制系列产品，特别是在高带宽情况下的网络流量处理方面具备一定的技术优势。经过几年的不懈努力，东华软件在此方面已经成为国内的领先企业，建立了完备的产品体系。网络流量分析、网络流量管控等产品在电信运营商、大型企业、国家政务网等行业拥有大量的成功案例。东华软件股份公司已为数千个用户提供了优秀的信息系统解决方案，涵盖多种应用与技术平台，用户遍布电信、电力、政府、交通、国防、医疗、金融、科研、煤炭、石化、石油、保险及制造等行业。