多核分布式 NetStream 技术白皮书

关键词：NetStream、NDE、NDA、NSC、流量统计、计费、ToS、NetFlow、多核、分布式

摘　要：本文介绍了 H3C 公司的多核分布式 NetStream 技术。NetStream 技术是一种基于网络流信息统计与发布的技术，它提供了一套对网络中通信量和资源使用情况进行统计并将这些数据用于管理、分析和计费的方案。而 H3C 的 Netstream 全面支持多核 CPU 以及全分布式处理，大幅度提升了设备整机网流分析处理的能力，为用户提供了一种经济的大容量 NetStream 解决方案。

缩略语：

缩略语	英文全名	中文解释
-	-	-

目 录

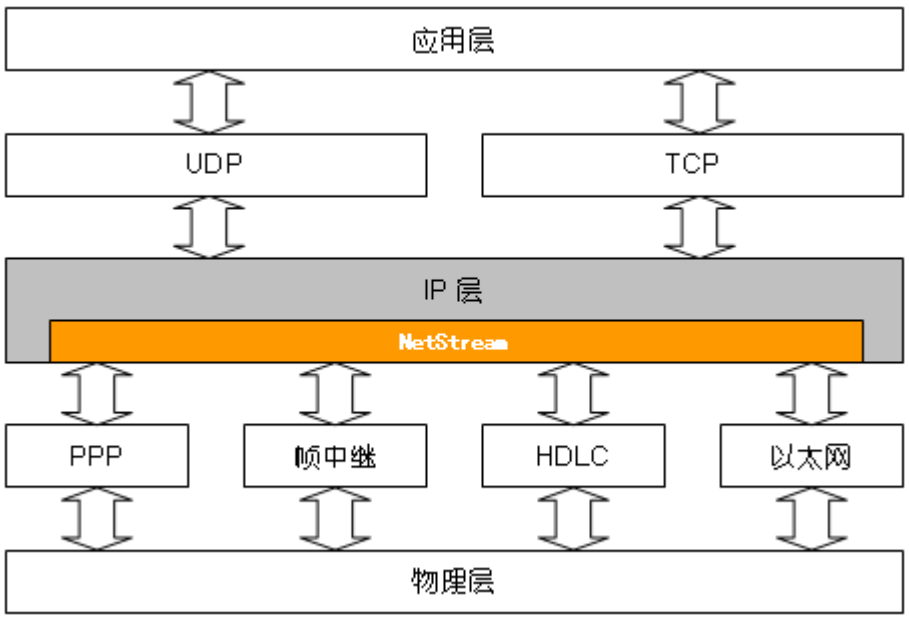
1 NetStream 技术介绍

1.1 概述

随着网络设备软硬件技术的快速发展，用户获得了更高的网络带宽，而网络支持的业务和应用也呈爆炸式的增长。为了对网络进行更细致的管理和计费，用户对流量统计分析提出了更高的要求。H3C 的 NetStream 技术是一种基于网络流信息统计与发布的技术，通过对网络中的通信量和资源使用情况进行统计和发布，为网络管理人员提供了解网络访问通讯量详细信息的途径。NetStream 输出的数据有许多用途：网络管理和规划、企业记账和分部门的计费、ISP 编制帐单、数据储备以及其他用于商业目的的数据采集。

如下图所示：NetStream 工作在 IP 层。网络设备在进行 IP 报文转发过程中，对于数据链路层（如以太网、PPP、HDLC、帧中继等）送入 IP 层的报文以及 IP 层发到数据链路层的报文，以网络流的形式进行分类、统计，并将统计结果以特定格式的 UDP 报文发送到网络上的数据采集器、分析器完成最终的统计、分析，输出相关的报表。

图1 NetStream 在网络设备中的逻辑位置



1.2 相关术语

- 网络流（NetStream）：一组特征相同的报文的集合。
- NDE（NetStream Data Exportor）：NetStream 原始数据加工输出设备。
- NSC（NetStream Collector）：网络流数据收集器。
- NDA（NetStream Data Analyzer）：网络流数据分析器
- AS（Autonomous System）：自治系统，在 BGP 路由协议中应用。

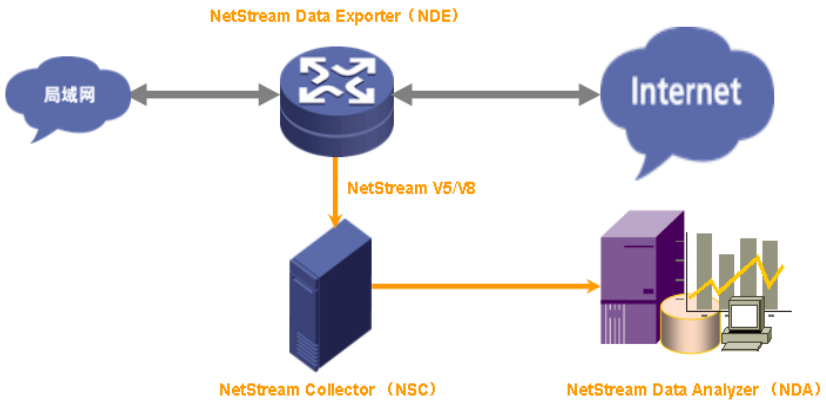
- TOS（Type of Service）：服务类型，在 IP 报文头中的标志，用来进行流量控制。

1.3 技术说明

NetStream 对流的定义为在一段时间内由源到目的之间的若干数据包，可以由 IP 地址和 TCP 或 UDP 协议端口号确定其特征，此外 NetStream 还使用 IP 协议类型、ToS 和输入（或输出）接口来唯一标识一条流，即构成如下的七元组：

- 源 IP 地址
- 目的 IP 地址
- 源端口号
- 目的端口号
- IP 协议类型
- 服务类型（ToS）
- 输入/输出接口

图2 NetStream 数据采集与分析



如上图所示，一个典型的 NetStream 系统由 NDE、NSC 和 NDA 三部分组成。

(1) NDE（NetStream Data Exporter）

通常为路由器。NDE 使用流缓存来维护流的相关信息和处理规则，根据流的第一个报文建立一个处理规则，后续利用该规则对后续网络流进行分析处理，提取符合条件的流统计信息，并将统计信息输出给 NDC 设备。输出前也可对数据进行一些处理，比如聚合。

- NSC（NetStream Collector）

通常为运行于 Unix 或者 Windows 上的一个应用程序。负责解析来自 NDE 的报文，把统计数据收集到数据库中，可供 NDA 进行解析。NSC 可以采集多个 NDE 设备输出的数据，对数据进行进一步的过滤和聚合。

- NDA（NetStream Data Analyzer）

NDA 是一个网络流量分析工具，它从 NDC 中提取统计数据，进行进一步的加工处理，生成报表，为各种业务提供依据（比如流量计费、网络规划，攻击监测），通常，NDA 具有图形化用户界面，使用户可以方便地获取、显示和分析从 NetStream Collector 收集的数据。

2 多核分布式 NetStream 技术特性

2.1 全分布式 NetStream 特性

业界目前在分布式设备实现 NetStream 特性（NDE）时多采用插入松耦合的 NetStream 业务板方式。各个 IO 板上根据配置好的规则，将特定的流量镜像到 NetStream 板上完成流信息的预处理（包括分类、聚合），然后再将统计信息发送给 NSC 以及 NDA 进行后续处理。

表1 NetStream 两种实现方式对比

	NetStream 业务板集中处理	H3C SR6608 NetStream 多核分布式处理
性能	分布式设备往往吞吐量很大，流量镜像到固定业务板做集中式 NetStream 处理，容易在 NetStream 业务板形成性能瓶颈	接口板三层接口上使能 NetStream，流量分析直接在单板完成，不存在性能瓶颈，并且可以大幅度提高了整机 NetStream 的处理能力（全分布式 NetStream）
部署成本	用户如需使用 NetStream，必须购买专用的业务板， 部署成本较高并且占用了宝贵的槽位资源	无需另外配置 NetStream 板，节省了用户的投资

从上面的对比表格可以看出：传统集中式 NetStream 业务板经过从功能实现上来说是没有问题的，但还是存在性能以及部署成本上的缺陷。而 H3C 公司在其 SR6608 产品上实现的全分布式多核 NetStream 方案完美的解决了上述问题。

2.2 多核 NetStream 特性

对于传统单 CPU 网络设备，报文处理和命令配置全部由单个 CPU 完成，受到 CPU 处理能力的限制，产品的性能很难有大幅度的提升，并且业务支持的越多，对设备控制层面的功能影响越大（配置操作响应缓慢等）。另外单凭提升 CPU 主频提高性能因为技术上存在障碍也变得愈发艰难。ASIC、NP 技术虽然能够提供很高的处理性能，但是通用性差、开发周期长且不适于复杂业务的处理，不能满足多业务灵活处理以及快速推出业务的需要。

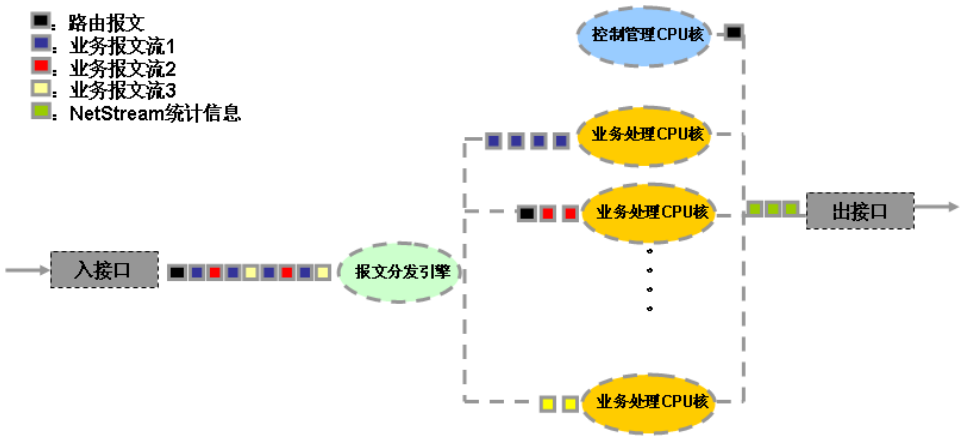
多核 CPU 的出现提供了解决上述问题的思路。

多核 CPU 可以看作是将多个通用的 CPU 以及一些功能部件集成到一块芯片中形成的一个 SOC（System On Chip），这些 CPU 之间以及 CPU 与集成到芯片上的其他部件间通过高速的内部互联技术进行通信，打破了以往多 CPU 系统中 CPU 之间以及 CPU 与系统其他部件间通信的性能瓶颈，并且能够并行处理任务，从而使系统性能得到大幅度提升。

在家庭娱乐以及商用 PC 机型上，Intel 以及 AMD 公司的双核 CPU 已经得到了普遍的应用，在任务并行处理上发挥了强大的效能，另外性能更强的四核 CPU 也已经面世并且得到应用。

在专业通信设备应用中，多核 CPU 也开始崭露头角。H3C 公司率先在业界推出了商用多核路由器设备——SR6600 路由器。在 SR6600 路由器上使用了专业多核多线程 CPU 进行网络报文转发和业务的并行高速处理。该多核 CPU 内部包含了 8 个 CPU 核，每个核又包含 4 个硬件线程（Thread），多个 thread 共享流水线和一级 CACHE，每个 thread 分别拥有各自的硬件寄存器。整个系统运行起来就如同有 32 个 CPU（8×4）在并行处理业务。

图3 H3C 多核 NetStream 软件架构



上图为 H3C 公司基于多核 CPU 搭建的多核 NetStream 软件架构示意图。

该架构的特点可以总结如下：

- 高效的负载均衡：硬件报文分发引擎按照一定的算法将报文按流或者业务类型均匀高速分发到业务 CPU 核上，有效的避免了业务 CPU 核负荷超载或者不足。
- 强大的并行网流分析：基于多核软件架构的 NetStream NDE 并行运行在多个业务处理 CPU 核上，对网络流量按照规则进行分析、聚合，输出相关的统计信息给 NSC 用以后续处理。由于业务是并行处理，从而极大提升了设备的 NetStream 处理能力。
- 完美地实现了控制与业务的隔离：控制平面和数据平面运行在不同的 CPU 核上，互不干扰，从而最大程度地保证了系统的管理功能稳定以及业务的运行稳定。

2.3 对入接口和出接口分别独立进行统计

Netstream 流统计时按照出入接口独立统计。

对于出接口：由出接口、源 IP 地址、目的 IP 地址、源端口号、目的端口号、协议号和 TOS 组成的七元组确定一条流（对于路由负载分担的情况，流统计信息中的下一跳和出接口从第一个负载分担路由项获取）；

对于入接口：由入接口、源 IP 地址、目的 IP 地址、源端口号、目的端口号、协议号和 TOS 组成的七元组确定一条流。

2.4 统计信息的老化

在实际网络环境中，可能在很短的时间内产生大量流，而 NDE（路由器设备）的内存容量又是有限的，这就需要根据一定的算法把一部分流从 NDE 的 NetStream 缓冲区中删除，释放内存，同时以 UDP 数据包的形式向 NSC 输出关于这条流的统计信息，这称为统计信息的老化（aging）。

H3C 的 NetStream 提供四种老化机制：按时老化、TCP 的 FIN 和 RST 报文触发老化、统计字节数溢出老化、命令行强制老化。下面将分别加以介绍。

2.4.1 按时老化

通常，网络中的数据流量具有突发性的特点。对于 NDE 设备，第一个 60 秒钟内可能有 10 万个属于同一条流的数据包通过，而在第二个 60 秒钟内，此类数据包可能一个也没有。

针对这种情况，NetStream 有以下两种按时老化的机制：

- 按不活跃时间老化：对从最后一个报文流过到当前的时间超过 inactive timeout 的流进行老化；
- 按活跃时间老化：对从第一个报文流过到当前的时间超过 active timeout 的流进行老化。

H3C 的 NetStream 对于两种按时老化均支持，老化时长也可以配置。

2.4.2 TCP 的 FIN 或 RST 触发老化

对于 TCP 连接，当有标志为 FIN 或 RST 的报文发送时，表示一次会话结束。因此当一条已经存在的 TCP 协议 NetStream 流中流过一条标志为 FIN 或 RST 的报文时，将立即把相应的 NetStream 流老化掉，并输出统计信息给 NSC。但是假如一条流的第一个报文就是 TCP 的 FIN 或 RST 报文，则根据正常流程创建一条新流，然后继续执行，不进行老化。

2.4.3 统计字节数溢出老化

NetStream 流缓存区中的流需要记录流过的总报文字节数，对于一个典型的 32 位系统的中整数的最大上限约为 4G，当字节数超过变量类型的最大上限时，如果继续进行累加统计将会发生溢出进而导致统计发生错误。所以 H3C NetStream 功能在工作的时候，如果检测到某条流的字节数统计将要超过限制（达到 3.9G 字节）时，就立即把这条流老化，并且输出统计信息给 NSC。

2.4.4 命令行强行老化

允许用户通过控制台输入相关命令，将缓存中的流全部老化，并输出统计信息给 NSC。

2.5 NetStream 统计信息输出

NDE 缓存流信息老化的同时，流统计信息将以如下图所示的 UDP 数据报发往 NSC。

通过支持 sequence number，NSC 可以检测统计报文丢失情况。

图4 NetStream 报文格式

目前 NetStream 业界流行的格式有三种：V5/V8/V9。三种格式报文优缺点对比如下表：

表2 NetStream 报文格式对比

版本	优点	缺点
V5	• 输出的字段比较丰富，可以把聚合前的流记录的所有字段都输出给 NSC • NDE 设备负荷较小	• 报文格式固定且不可扩展 • 数据量大，NSC 无法长期保存，NSC 和 NDA 分析压力大
V8	• 按一定规则进行聚合，数据量相对较小 • 承载内容略为简单，适合特定分析 • 可以增加新的聚合方式	• 报文格式固定且不可扩展 • NDE 设备完成聚合工作，负荷较重 • 如果增加新的聚合方式，需要 NDE 和 NSC 同时升级版本
V9	基于模板 • 允许单独输出需要的域统计信息，减少了输出流的数据量进而减少了 NDE、NSC 可能的内存以及带宽开销 • 不需要改变输出报文格式即可在输出记录中增加新的域，因为是以模板形式输出所以即使 NSC 无法理解新增的域的真正语意，它仍然可以解释流记录 • 输出方式灵活，既可输出聚合前的流纪录，又可输出聚合后的流纪录	

说明

H3C 的多核分布式 NetStream 目前支持 V5 格式和 V8 格式的统计信息报文。V9 格式的统计信息报文即将支持。

1. V5 格式的统计信息报文

V5 格式的 NetStream 报文中包含每条流的原始信息，它由一个报文头和若干条报文记录组成，这些记录都对应着一条被老化的 NetStream 流。

对于 NetStream 出统计和入统计，分别生成各自独立的 UDP 版本 5 报文，两种报文格式一样，但是带有区别出统计和入统计的标志位。

2. V8 格式的统计信息报文

NetStream V5 采用的输出报文格式，在网络流量较大的情况下，会产生大量的 NetStream 输出报文，为降低这种情况造成的影响，NetStream V8 格式将流中原始信息按照一定的规则进行分类、合并后生成聚合的信息，再发送出去。对于 NetStream 出统计和入统计，分别生成各自独立的 UDP V8 报文，两种报文格式一样，但是带有区别出统计和入统计的标志位。

Netstream 聚合的好处显而易见：

- 减少了 NetStream 输出的数据量，从而减少了路由器设备和网管设备之间传输统计信息的带宽占用。
- 减少了 NetStream Collector 设备的工作量，降低了对 NetStream Collector 设备性能的要求。

目前，H3C 多核分布式 NetStream 支持以下 5 种聚合：

表3 H3C 多核分布式 NetStream 支持的聚合方式

聚合类型	聚合分类原则
自治系统聚合：as	根据 NetStream 流的源自治系统号、目的自治系统号、输入接口索引、输出接口索引 4 个关键值对流分类
协议－端口聚合：protocol-port	根据 NetStream 流的协议号、源端口、目的端口 3 个关键值对流分类
源前缀聚合：source-prefix	根据 NetStream 流的源自治系统号、源掩码长度、源前缀、输入接口索引 4 个关键值对流分类
目的前缀聚合：destination-prefix	根据 NetStream 流的目的自治系统号、目的掩码长度、目的前缀、输出接口索引 4 个关键值对流分类
源和目的前缀聚合：prefix	根据 NetStream 流的源自治系统号、目的自治系统号、源掩码长度、目的掩码长度，源前缀、目的前缀、输入接口索引、输出接口索引 8 个关键值对流分类

3. V9 格式的统计信息报文

如 表 3分析，V5/V8 格式的报文灵活性还有所欠缺，V9 采用模板方式解决了这一难题。

NetStream V9 采用两种类型的数据与 NSC/NDA 一起配合完成流的统计：一种是统计数据，一种是选项数据。

2.6 兼容性

H3C NetStream 的版本 5 格式对入接口的统计信息输出与 CISCO NetFlow 的版本 5 格式统计信息输出完全兼容。

H3C NetStream 的版本 8 格式对入接口的统计信息输出与 CISCO NetFlow 的版本 8 格式统计信息输出完全兼容。

3 NetStream 管理应用工具

H3C 公司提供了强大的 NetStream 管理应用工具 Xlog。

XLog 是可扩展的网络日志审计系统（Extendable Network Log Audit System）的英文简称，工具包括 NSC、NDA 两部分。该工具与 H3C 的路由器、交换机产品 NetStream 特性配合，可提供如下功能：

- 构建一个可扩展的和分布式的 NetStream 数据流收集和分析系统。
- 收集和存储多个 NDE 的数据，对数据进行过滤、聚合、存储入库。
- 进一步对数据进行分析产生流量报表。采用基于 Web 形式访问，提供直观和图形化的管理界面，所有数据输出都以友好的形式直接在 Web 页面中显示。

通过这个工具，可以为客户提供了一种可靠的、便利的网络流量分析解决方案，可以帮助网络管理人员了解企业内部网络运行状况，及时发现并解决网络中的性能瓶颈问题、网络异常现象，方便网络管理员及时解决网络异常问题，并能作为用户进行网络规划、网络优化、故障诊断等工作的参考。

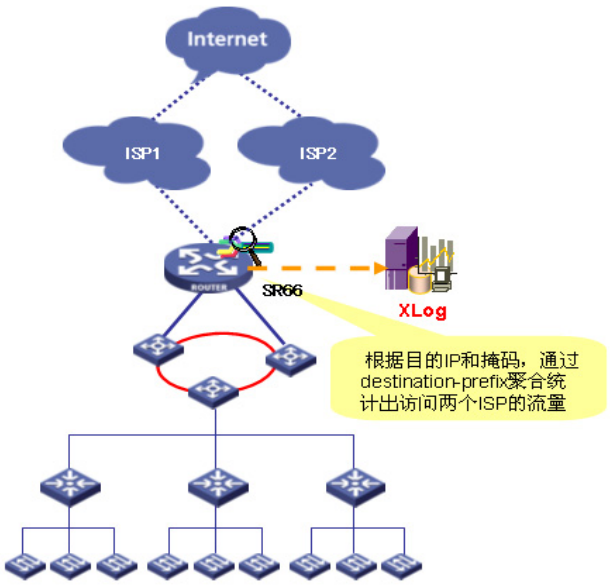
关于 Xlog 更详细的介绍，请参考该产品相关的用户手册。

4 NetStream 技术的应用

NetStream 技术以流为基础，实现了详细的网络信息收集。NetStream 技术由 H3C 的路由器、交换机提供，配合 H3C 的 NetStream 管理应用工具 Xlog，能够支持多种用户应用。

4.1 计费

图5 NetStream 用于 ISP 间流量分帐计费

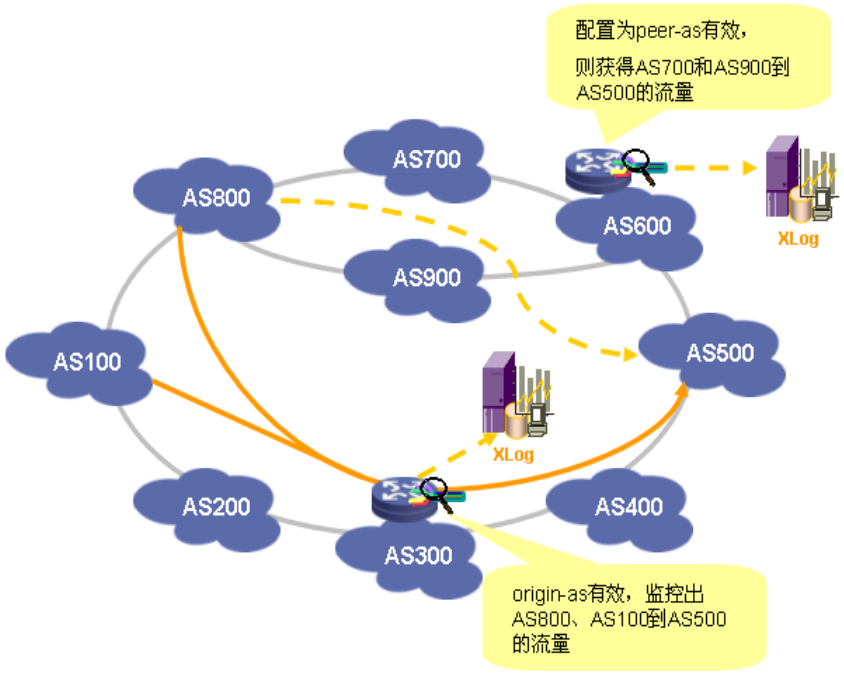


NetStream 为基于资源（如线路、带宽、时段等）占用情况的计费提供了精细的数据，这些数据包括 IP 地址、报文数、字节数、时间、TOS 和应用类型等。Internet 服务提供商可以利用这些信息来实行灵活的计费策略，如基于时间、带宽、应用、服务质量等。企业客户可以使用这些信息计算部门费用或分配成本，以便有效利用资源。

如上图所示的网络经 H3C SR66 路由器通过双 ISP 接入 Internet，在 SR66 上配置 NetStream 后，ISP 可以根据 NetStream 给出的流量统计完成分帐计费。

4.2 网络规划

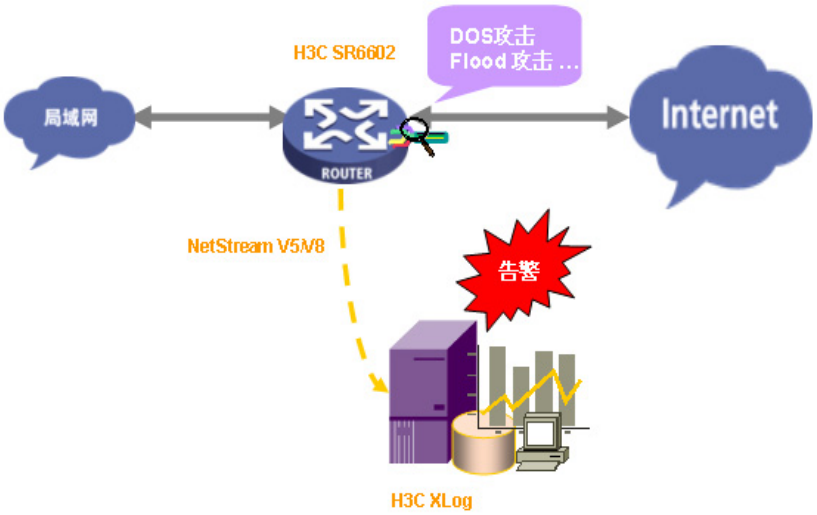
图6 NetStream 用于网络规划



如上图所示，NetStream 可以为网络管理工具提供关键信息，比如各个 AS 域之间的网络流量情况，以便优化网络设计和规划，实现以最小的网络运营成本达到最佳的网络性能和可靠性。

4.3 网络监控

图7 NetStream 用于网络监控



如上图所示，大多数局域网都会最终通过路由器连接到 Internet 网络。通过路由器出口部署 NetStream，对 Internet 出口进行实时的流量监控，可以分析各种业务占用出口带宽的情况，监视非工作需要的 Internet 访问，并且在攻击发生的时候及时在 Xlog 网管上发出告警，以便网管人员分析解决方案，排除故障。

4.4 应用监控和分析

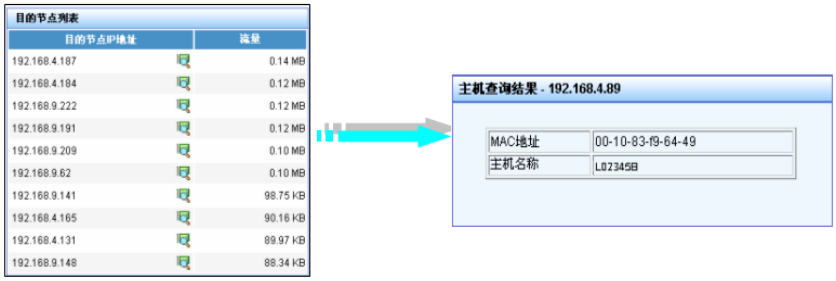
通过 NetStream 技术，可以获得详细的网络应用信息。例如，网络管理员可以查看 Web、文件传输协议(FTP)、Telnet 和其它常用的 TCP/IP 应用所占通信量的百分比。Internet 内容和服务提供商可以根据这些信息来规划和分配网络和应用资源以满足用户需求。

如下图所示，H3C 的 Xlog 解析 NetStream 信息，可以进行基于应用的统计，统计结果通过图形化界面显示，一目了然。目前 Xlog 支持近 300 种应用，并且用户还可根据自己网络的情况自行添加应用。

图8 NetStream 用于应用监控和分析

4.5 用户监控和分析

图9 NetStream 用于用户监控和分析



通过 NetStream 技术使得网络管理者可以轻松获取用户使用网络和应用资源的详细情况，进而用于高效地规划以及分配网络资源，并保障网络的安全运行。上图就是 H3C Xlog 通过分析 NetStream 信息获取到的特定用户的网络流量统计结果。

5 总结

在数据网络运营管理显得愈发重要的今天，多核分布式 NetStream 技术的特点有效地解决了以往流量分析方案存在的性能瓶颈和部署成本昂贵的问题，而采用这一技术的 H3C SR6608 路由器能够帮助网络运营者全面了解大容量网络的流量细节，从而使其网络运行井然有序，将有限的带宽利用到有价值的应用上，发挥最大的效力。