IPFIX 网络流量分析技术研究与系统设计

刘克难，赵慧娟，魏俊超
(西南交通大学 信息科学与技术学院，四川 成都 610031)

摘 要：流量分析技术是网络管理的重要组成部分，为网络管理提供数据支撑。描述了基于流的 IP 流导出标准 IPFIX(IP Flow Information eXport)的工作原理，分析了 IPFIX 流形成与处理的整个过程，并对流量分析系统结构进行研究设计。针对分析器可能产生的性能瓶颈问题，提出两点解决方案，并描述了收集器负载均衡的实现方法。

关键词：流量分析；分析器；收集器；负载均衡
中图分类号：TP393 **文献标识码**：A **文章编号**：1674-7720(2011)04-0011-03

Research and system design of IPFIX flow analysis technology

Liu Kenan, Zhao Huijuan, Wei Junchao
(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: The technology of flow analysis plays an important part in network management, and it provides relational data to support network management. The flow analysis which based on flow technology is used widely because of its advantages. This thesis describes the operational principle of IPFIX, which is an IP flow export standard based on flow technology. The procedure of flow generation and process are illustrated, and the architecture of flow analysis system is researched in this thesis. Two solutions are provided to avoid the bottleneck of the analyzer performances, and load balance methods of the collectors are implemented in this thesis.

Key words: flow analysis; analyzer; collector; load balance

网络管理技术是联网系统中一个重要组成部分。它为网络高效稳定地运行提供了重要保证。流量分析技术作为网络管理的重要组成部分，其发展也十分迅速。

流量数据的收集和分析有多种手段，常用的有基于 SNMP 协议的接口统计方式、RMON 方式、数据流量探针等[1-2]，这些方式各有特色，适用于网络中不同环境的流量分析工作。但由于其统计流量方式不灵活或成本高等原因，无法满足对网络进行更细致管理的要求，需要一种新技术来更好地支持网络流量统计与分析。于是，专门用于网络流数据收集的协议产生了，如 NetFlow、SFlow、NetStream 等基于 Flow 的技术，由于其实施成本低、操作灵活、扩展性好，并且不需要额外硬件支持的显著特点，使得这些技术很快得到广泛应用，特别是思科的 NetFlow 协议。

然而，在实际应用中，由于各个大厂商各自拥有自己的 Flow 格式和流量数据输出系统，使得彼此之间缺乏兼容性，无法满足大规模异构网络中应用的需要。因此，无论是在学术界还是工业界，都有必要建立一套输出网络流信息的标准。于是 IETF 在应用广泛的 NetFlow V9[3]的基础上制定了 IPFIX(IP Flow Information eXport)技术标准。IPFIX 在 NetFlow 的基础上对安全、数据存储等方面进行了改进与扩充。提供了流量分析系统与网络设备之间数据采集与传输的规范，定义了数据交互的标准格式，保证了对采集目标网元设备良好的兼容性。

1 IPFIX 技术原理

1.1 概念

观察点[3]：指网络中用来观察和获取 IP 数据包的位置，例如以太网 LAN 中的一个或一组接口设备(路由器或交换机等)。

流(Flow)：在一定时间间隔内，通过一个观察点并具有相同数据包属性和操作动作的数据包的集合。流的属性包括测量属性(总的字节数、总包数等)与特征属性(IP、端口等)。一个数据包被归于某流，当且仅当其完全满足该流的所有特征属性(键值)。

模板(Template)：定义 Flow 记录中各个属性字段的结构和语义。收集进程根据模板来解析数据流，当需要

支持某种新格式数据流时，只需变换模板，不需要对程序进行修改，使得 IPFIX 具备良好的可扩展性。

可选模板：可选流记录提供收集器需要的一些额外信息，如流的键值、采样、聚合策略等。可选模板定义了可选流记录中的各个属性字段的结构和语义。

1.2 工作原理

IPFIX 以一种灵活的、可扩展的输出流格式为用户提供更加细致的流量信息。下面从 Flow 记录的形成、导出、收集的过程来详细分析 IPFIX 的工作原理[4]，如图 1 所示。

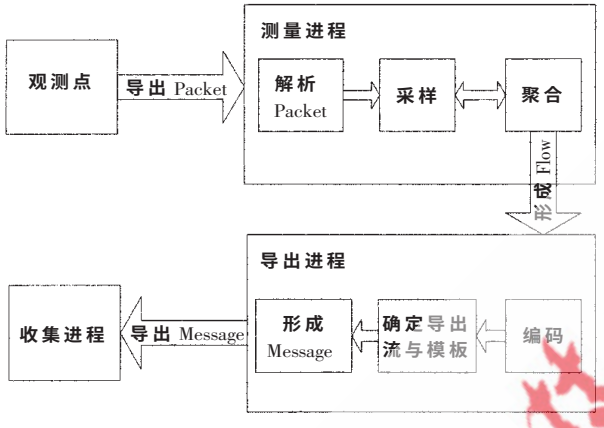

图 1 IPFIX 流处理过程

首先测量进程(Meter)，从观测点观测到数据包(Packet)，解析数据包首部、添加时间戳，为了减少传输流的数量，或只产生满足某种需要的流记录，可以采用采样与聚合技术，采样与聚合的规则要明确定义，并由导出进程以可选模板与可选信息记录的形式发送到收集器。不能将使用不同规则的聚合策略采样得到的数据混在一起，当规则有变化时，应立即同步处理。测量进程主要责任是产生新的流记录、更新或删除旧的流记录、定义并计算流的寿命、计算统计值。在此过程中发生超时、过载等任何异常都会记录日志，以方便跟踪分析。

导出进程(Exporter)，产生控制信息(模板、可选模板)，将控制信息与数据流信息一同装载到 IPFIX 消息(Message)中，发送到收集进程。收集进程根据控制信息完成相应的数据流的解析。IPFIX 协议中，导出进程只导出老化后的流，老化规则由测量进程确定，一般满足下列条件之一的流被视为老化：

(1)该类型数据流传送完毕；

(2)在一个时间段内某个数据流处于不活跃的状态，这个时间段可以由输出器进行配置；

(3)对于一个长时间活跃的数据流，收集器将按照规则输出数据流记录，并强制将该数据流老化，这个“时间段”应该在输出器上进行配置；

(4)在导出进程内部条件受限的情况下，数据流也会被强制进行老化。例如：计数器归零或内存不足。

为了安全考虑，可以将信息进一步编码加密。同样，导出进程也对错误进行记录日志，并统计丢包个数等数据。

导出进程可以将同样的数据发送到多个收集进程，多个收集进程将这些数据汇总，汇总时要注意处理好重复的数据。这种冗余数据的做法增强了流数据的完整性，但加大了负载，只有在对数据丢失相当敏感的应用中可以采用。

导出进程向不同的收集进程传送数据可以采用不同的协议，同 NetFlow 默认使用 UDP 不同，IPFIX 默认的是 SCTP[5]。当收集器收到一个不正常的 IPFIX 消息时，会话会被重新发起，并丢弃本次会话的所有模板、停止解析、登记日志。相同的模板与可选模板只传送一次。收集器必须存储下来。如果收集进程先收到数据记录，则要等待模板，等待时间是可配置的。IPFIX 消息的序列号在首部，它随着 IPFIX 消息中的数据流记录的增加而增加。通过这个序列号，收集进程可以知晓是否有乱序的、丢失的或重复的数据，以跟踪处理。如 NetFlow 一样，IPFIX 的传输协议也可以使用 UDP 或 TCP，使用不同的协议，导出进程与收集进程间的传输过程管理也不一样，可参考 RFC5101。

2 流量分析系统结构

根据 IPFIX 的工作原理可知，其包含三个重要组成部分，即测量进程、导出进程、收集进程。参照其各自的功能，将流量分析系统按功能分成：探测器、收集器、分析器三个主要部分进行分布式部署(如图 2)。其中，分析器接收至少一个收集器的数据，收集器接收至少一个探测器的数据，探测器可同时向多个收集器发送数据。

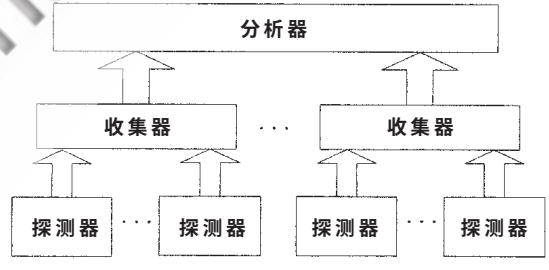

图 2 流量分析系统结构图

(1)探测器的工作对应于图 1 中的测量进程与导出进程部分，负责 IPFIX 流的形成与处理过程，最终形成 IPFIX 消息发送给收集器。它主动采集数据包，透明地穿越网络结构，不会修改任何报文。除非进行了配置，否则它不会放过任何报文。工作效率非常高，其输出数据的总量大约是网络设备之间交换量的 1.5%[6]；

(2)收集器对应图 1 中的收集进程，具体工作细节见图 3，它接受 IPFIX 消息，解析并形成压缩文件，定期将文件传送到分析器；

(3)分析器负责计算与统计分析，如图 3 所示，它接收 IPFIX 文件，并将信息存储到数据库中，按用户的需求，对数据库中的数据进行计算，形成报表呈现给用户。

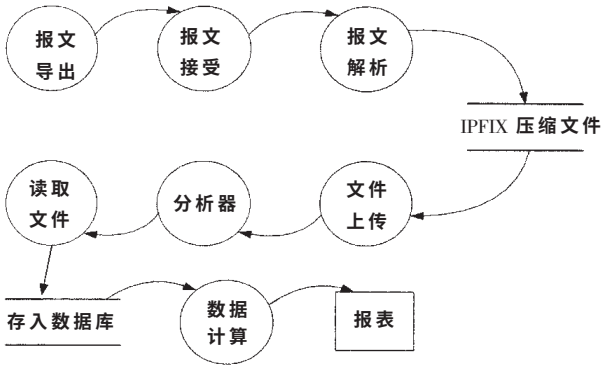

图 3 IPFIX 消息处理数据流图

总之，流量分析系统的工作是把网络中各种应用数据包进行清洗、筛选、标准化处理，形成属性明确的流，再对这海量的流记录进行数据挖掘的过程。挖掘出用户关注的信息，以清晰直观的方式展现给用户，为用户对网络管理或其他应用分析提供可靠的数据支持。

3 结构分析与优化

如图 2 的系统结构，实践中发现，流记录的海量数据全部放到分析器进行计算分析，分析器处理能力有限，产生了瓶颈，影响了整个系统的性能。

本文给出了两点改进方案：计算下移和收集器负载均衡。

(1)计算下移

收集器只是简单地将数据解析并存储，把所有的计算任务都交给分析器。经常出现的情况是收集器资源利用不充分，分析器过载。故考虑在收集器将数据发送给分析器之前，需对数据进行一些预处理工作，这样就能缓解分析器的工作负载。如主机流量排行、应用流汇总数据，这些常用的分析结果可以在收集器上进行排序与计算后，再将数据传送给分析器，必要时，可以将预计算的结果存放在另一文件中与 IPFIX 文件一同上传分析器。分析器从这一文件直接得到预计算后的数据，而不必再对元数据进行计算，从而减少了分析器的负载，提高了分析效率。

(2)收集器负载均衡

多个收集器同时工作时，有的收集器可能很忙碌，甚至达到性能极限，而此时，其他收集器则处于空闲状态，形成了资源的浪费，降低了工作效率。因此对收集器进行负载均衡是必要的，尤其对于计算下移的系统。负载均衡，即根据收集器的几个重要的性能指标来衡量收集器的资源使用情况，并根据使用情况，动态调度各收集器的工作任务。

本系统在分析器中设计了一个均衡器来实现均衡调度。均衡器维护一张收集器资源使用情况实时状态表。具体采用基于负反馈机制的动态负载均衡算法[7]，该算法考虑每个收集器的实时负载情况，不断调整任务的分配比例，避免有些收集器超载时依然收到大量 IPFIX 消息，从而提高系统整体工作效率，使整个系统成为一个智能的有机体。在系统内，分析器承担管理角色，其均衡器负责监听各个收集器的负载信息，收集器定时向均衡器报告自身的负载状况。分析器根据均衡器的实时信息，对各个收集器的工作量进行调整。对于超载的收集器，削减其工作量，并取消其对 IPFIX 消息的预处理工作。

收集器负载值的计算主要根据其 CPU、内存以及磁盘使用状况得出。假设某系统内有一组收集器 C_i $(i=1,\cdots,n)$，$L(C_i)$ 表示收集器的 C_i 负载值，$L_X(C_i)$ 表示收集器 C_i 的性能指标 X 当前负载值。给 $L_X(C_i)$ 设定一个系数 R_i $(i=1,\cdots,n)$，用来表示各个性能参数的重要程度。

收集器 C_i 的负载值公式可描述为：

$$\begin{cases} L(C_i)=R_1L_{CPU}(C_i)+R_2L_{memory}(C_i)+R_3L_{IO}(C_i) \\ \sum_{i=1}^{3} R_i=1 \end{cases}$$

收集器主要工作是解析 IPFIX 消息，以观测点为单位，将解析后的数据写入文件，并将文件周期性地发送给分析器，发送成功的文件做本地删除。针对其特点与实际经验，可取系数{0.4，0.4，0.2}。若当前的系数不能很好地反映实际应用的负载情况，可以对系数及时地进行修正，直到找到贴近当前应用的一组系数。

关于计算 $L(C_i)$ 的周期设置，虽然较短的周期可以更准确及时地反映各个收集器的负载情况，但是频繁的计算会给均衡器和收集器带来负担，也增加了不必要的网络负荷。另外，频繁计算会导致均衡器反映出的负载信息出现抖动，均衡器无法准确捕捉各收集器实际的负载变化趋势。虽然可用数据拟合等技术来处理数据，使负载信息表现为平滑曲线，但如此额外增加了收集器的负担，故应避开抖动，适当增长计算周期。

每个参数都设置有阈值，如果某收集器的某一性能指标超过阈值，则立即减少该收集器的工作量，即应将向该收集器发送的 IPFIX 消息发送到其他空闲的收集器。如果系统中所有的收集器长期处于过载状态，则需要向系统中添加收集器，或调整采样与聚合策略来减少数据流。若所有收集器的使用率长期处于低状态，则可以削减收集器，减少不必要的维护。

IPFIX 作为 IP 流量导出的标准，继承并完善了 NetFlow。它提供了多种字段类型，可针对具体的应用自定义流的键值、模版、采样、聚合机制以及传输方式、编码规则等，能够详细描述流量的各类特征，为快速、准确地进行流量分析提供了坚实的数据基础。基于模板的灵活数据格式，使得其具备良好的可扩展性。本文根据 IPFIX 的工作原理，设计出流量分析系统的结构，并分析了可能出现的性能瓶颈，提出了两种解决此瓶颈的方案。随着网络应用的广泛与复杂，对流量分析的工作要求会进一步增高，需要更加高效合理的系统结构，如分析器分布式部署，并行分析将会成为未来研究的焦点。

参考文献

[1] 王珊，陈松，周明天.网络流量分析系统的设计与实现[J]. 计算机工程与应用，2009,45(10):86-88.

[2] 李兴国，费玲玲.基于 NetFlow 的流量分析技术研究[J]. 微计算机信息，2008(5-3):198-200.

[3] QUITTEK J, ZSEBY T, CLASIE B, et al. Requirements for IP flow information export(IPFIX)[EB/OL]. RFC3917, 2004(10).

[4] CLAISE B E. Cisco systems NetFlow services export version 9[EB/OL]. RFC3954, 2004(10).

[5] CLAISE B. Specification of the IP flow information export (IPFIX) protocol for the exchange of IP traffic flow information[EB/OL]. RFC5101, 2008(1).

[6] CLAISE B, WOLTER R. 网络管理：计费与性能管理策略[M].北京：人民邮电出版社，2009.

[7] 陈勇. 一种高效的分布式反馈流量负载均衡算法[J]. 计算机工程，2009,35(2):198-102.

(收稿日期：2010-10-31)

作者简介：

刘克难，男，1985 年生，研究生，主要研究方向：网络管理技术。

赵慧娟，女，1962 年生，教授，主要研究方向：网络管理，集群供应链协同方法，制造业信息化关键支撑技术，协同商务理论研究。