

# TCP 数据流超时研究

王远, 丁伟, 龚俭\*

(CERNET 华东(北)地区网络中心, 东南大学计算机科学与工程学院, 江苏 南京 210096)

**摘要:** K.C.Claffy 于 1994 年基于一个小时的主干数据分析, 提出了主干网数据流超时时间以 64 秒为宜的结论, 这个结论被广泛沿用至今。在此期间互联网发生了翻天覆地的变化, 这个结论是否仍然适用呢? 本文以 CERNET 某 2.5G 省网边界链路近期连续一小时报文数据构成的 trace 为对象, 采用 K.C.Claffy 相似的方法进行了统计分析, 得出的结论是: 目前的数据流超时时间间隔取值 16 秒时就可以达到理想的效果。这个结论将对所有与网络数据流有关的工程和研究具有借鉴意义。在实验数据量方面, 相对于 K.C.Claffy 的平均每秒 1118 个报文(下文简称 pps), 本次试验所采用的 TRACE 数据量为 650k pps, 约是前者的 581 倍。本文将详细介绍试验和结果产生的过程。

**关键词:** 网络数据流; 超时; 主干网; 网络测量

中图分类号: TP393.0

文献标识码: A

数据流(下文简称“流”),是指符合特定的流规范(flow specification)和流超时(flow timeout)约束的一系列数据包的集合,将数据包识别为流的过程一般被称为“组流”(flow profiling)。基于流(flow)的网络测量研究正被越来越多的研究者采用,它相对于传统的局限于数据包(packet)层次的研究在很多方面具有不可替代的优势[1][2][3]。

同一个数据包集合,采取不同的流规范和流超时,会产生不一样的组流结果。由于目前基于流的特定研究一般都具有明确的流规范,但在流超时方面却没有一定之规,因此对组流方法(flow profiling methodology)的研究较多的集中于流超时方面。K.C.Claffy 于 1994 年提出的“主干网数据流超时时间以 64 秒为宜”的结论,被业界广泛接受。但十几年来,网络规模和设备的发展日新月异,64 秒是否仍适合当前网络测量研究已成为疑问。本文的目的是,使用和 Claffy 相似的方法对近期 CERNET 主干网流量进行实验和分析,找到适合于当前网络状况的流超时值。本文的第 1 部分介绍了 Claffy 使用的实验方法和结论,第 2 部分比较了本文与 Claffy 在实验数据和评估标准方面的异同,第 3 部分展示并分析了实验结果,第 4 部分给出了结论和对未来研究的展望。

## 1 Claffy 的实验方法和结论

### 1.1 实验方法

Claffy[1][2]研究流超时的实验方法为:对同一段 trace,固定流规范(flow specification),采用不同的流超时值(flow timeout),分别组流(flow profiling),这样产生出的每个流集合之

---

基金项目:国家重点基础研究发展规划(973)课题(2003CB314804)

作者简介:王远(1983~),男,硕士研究生;丁伟(1962~),女,教授;龚俭(1957~),男,教授.

\*通讯作者: [ywang@njnet.edu.cn](mailto:ywang@njnet.edu.cn)

间的差异完全由超时值决定，从而可以将流集合在各方面的表现作为超时值优劣的评估标准，最后通过综合考虑其在所有评估标准上的表现，确定最优的流超时值。除了在评估标准方面有所改进之外，本文使用的实验方法与 Claffy 完全相同，采用的流规范为“主机对”（host pair），考查的传输层协议含 TCP、UDP 和 ICMP 三种，探索的流超时值（单位：秒）包括区间[2, 2048]中所有 2 的幂。具体的组流方法本文不再详述，读者可参见文献[1][2]。

## 1.2 评估标准

评估标准，是由具体的基于流的研究和应用决定的。Claffy 的评估标准主要包括以下几个方面：

### (1) 新流和活跃流个数

保存流信息的路由器相对于传统路由器有很多优势，如：报文转发快、拥塞控制合理等。对于这种路由器来说，新流多意味着建立和销毁流信息的操作频繁，这会大量消耗 CPU 资源，还将导致“颠簸”问题（thrashing）[1][2][3]；而活跃流多意味着需要大量的内存资源来维持流信息，而且查找流记录时间也将大大增加。所以，这种基于流的应用，其理想状态是新流和活跃流的个数都少，若无法实现理想，则应选择两者相对平衡的状态。

### (2) 流平均重建次数

很多基于流信息的研究，如：负载敏感路由（load-sensitive routing）[4]，需要长流（large flow）所占比例较大（一般认为，报文个数大于 10 的流为长流）[3]。所以，此类研究一般希望流平均重建次数较低，从而保证原本的长流不被截断（shortening）。在路由器方面，这个指标代表 CPU 的负担。

## 1.3 结论

Claffy 根据其实验结果和评估标准，得出以下关于流超时的结论[1][2]：在 16-128 秒之间的“主机对”流超时值似乎可以（seem to be）使路由器的处理任务和内存消耗达到平衡。其中 64 秒在各个指标上表现均衡，因此 Claffy 其后的所有实验都以 64 秒为超时值。

## 2 数据比较与新增评估标准

### 2.1 数据比较

表 1 和表 2 从多个方面展示和比较了 Claffy 和本文实验分别使用的 trace。其中，表 1 是 trace 的采集信息，表 2 是以一秒为单位进行统计的 trace 总体参数信息。

### 2.2 新增评估标准

本文除沿用 Claffy 的上述几个评估标准以外，还新增了一个由 Ryu[4]提出的“ $F_{hold}$  因子”指标。该指标定义如下： $F_{hold}(n) = D_{hold}(n) / D_{act}(n)$ ，其中  $D_{hold}(n)$  表示某个流的持续时间与其超时值之和， $D_{act}(n)$  表示该流的持续时间。此指标的意义在于：每一条流在处理完最后一个报文后，都不能立即结束，而必须在等待规定的超时时间后，才能被判定为结束，进而从内存中移除，所以

该值越小，说明这条流在内存里浪费的时间（流超时时间）与有用的时间（流持续时间）之比越小，也就是越期待的。“ $F_{hold}$  因子”类似于物理学概念中的“无用功比例”。 $F_{hold}(n) > 1.0$ ，最理想的情况是  $F_{hold}(n) = 1$ 。在路由器方面， $F_{hold}$  代表内存的负担。

对于由 N 条流组成的集合，其平均  $F_{hold}$  值计算公式为： $\bar{F}_{hold} = \left(\frac{1}{N} \sum_n F_{hold}^{-1}(n)\right)^{-1}$ 。因为单报文流的  $F_{hold}$  值为  $\infty$ ，所以这里使用  $F_{hold}^{-1}(n)$  来计算平均值。 $\bar{F}_{hold} > 1.0$ ，最理想的情况是  $\bar{F}_{hold} = 1$ 。

表 1. trace 采集信息  
Tab.1 collection information of trace

	采集点	网络层次	采集日期	采集时间	数据总量
Claffy	Urbana-Champaign FDDI into NSFNET	主干网	1993.3.23	14:11-15:11	4019k pkt
本文	CERNET 某 2.5G 省网边界链路	主干网	2005.11.10	14:00-15:00	2340m pkt

注：“pkt”表示报文（packet），两段 trace 的采集日均为工作日。

表 2. 以一秒为单位进行统计的 trace 总体参数信息  
Tab.2 population parameters for one-hour data sets measured by one-second intervals

	total	min	mean	95%	max	SD	total	min	mean	95%	max	SD
	kpkt	pkt/sec					MB	Kbytes/sec				
UC-NSF	4019	659	1118	1549	2355	228	1151	92	320	548	913	122
CERNET	2340k	483k	650k	684k	951k	36k	1490k	308k	414k	433k	449k	14k

### 3 实验结果与分析

#### 3.1 新流和活跃流个数

图 1 的上半部分显示了按秒排序，新流个数的中值和第 95 区间分别与流超时值的对应关系。图 1 的下半部分显示了按秒排序，活跃流个数的中值和最大值分别与流超时值的对应关系。如图所示，新流个数是超时值的减函数，而活跃流个数是超时值的增函数。因此，只能在两者之间选择一个平衡。注意到，在 16 秒处，新流的中值为 11932，活跃流的中值为 140374；64 秒处的新流中值虽然下降为 8142，但活跃流中值上升为 253722；再看 2 秒处，其活跃流中值为 84118，新流中值却上升到 26703。新流第 95 区间和活跃流最大值的情况也和其中值的表现类似。

假设路由器处理一个新流和一个活跃流的耗费分别是 N 和 A，单位时间内需要处理的新流和活跃流个数分别为 n 和 a，则总的耗费  $cost = Aa + Nn$ ，令  $F = N/A$ ，图 2 给出了在 F 的几个取值下，cost 与流超时值的对应关系。可以看到，在图 2 的这些情况下，16 秒是一个较好的平衡点。

#### 3.2 流平均重建次数

图 3 显示了流平均重建次数与流超时值的对应关系。在 2 秒处，平均每个流要被重建 7 次，而

在 16 秒处，该值首次下降到 3 次以下，为 2.56。这表明现有网络设备的 CPU 能够应付其重建任务。

### 3.3 $F_{hold}$ 因子

图 4 显示了各流超时值的与其  $F_{hold}$  值的对应关系。最优点，即  $F_{hold}$  最小处为 8 秒，其  $F_{hold}$  值为 3.62；16 秒的  $F_{hold}$  值为 4.16，比最优点仅高出 14.9%，在所有超时值中排名第三；而 64 秒的  $F_{hold}$  为 5.02，比最优点高出 38.7%，排名第六。

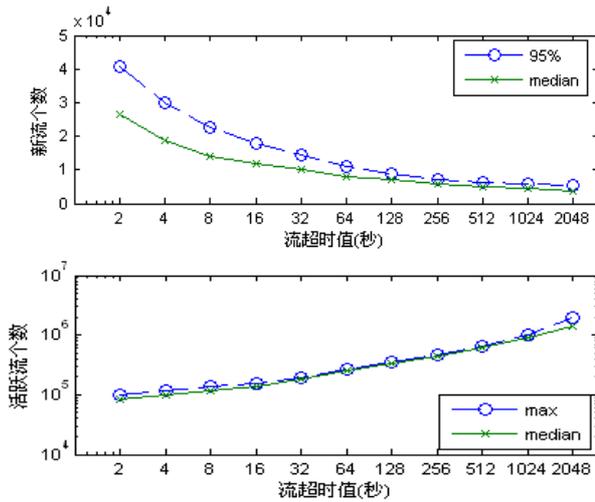


图 1. 按秒排序(上)新流个数的中值和第 95 区间分别与流超时值的对应关系. (下) 活跃流个数的中值和最大值分别与流超时值的对应关系.

Fig. 1 As a function of timeout value, (a) top: median and 95<sup>th</sup> percentile of new host pair flows per second (b) bottom: median and maximum of number of active flows per second.

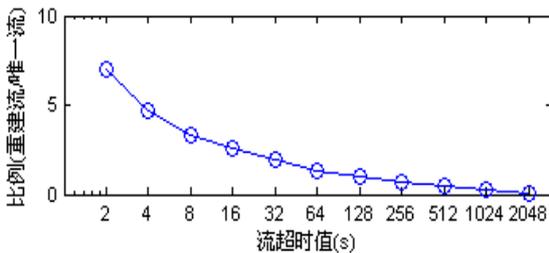


图 3. 流平均重建次数与流超时值的对应关系.

Fig. 3 ratio flows that were recreated within the flow timeout value to the 13M unique host pair flows.

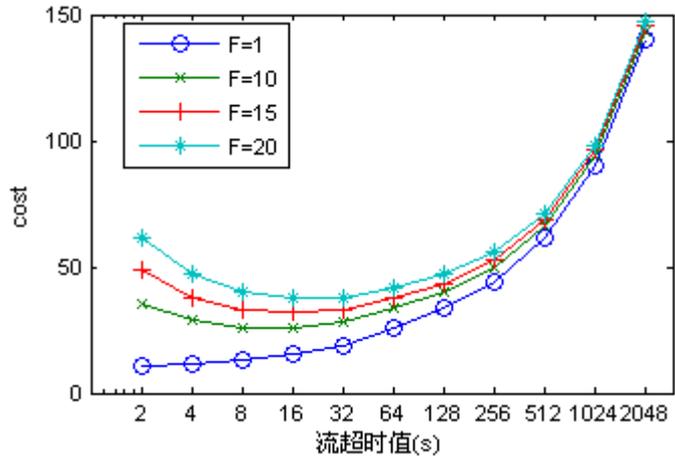


图 2. 在一些建立流耗费和保持流耗费比例下的，操作花费与流超时值的对应关系.

Fig. 2 Operating cost versus timeout value for various values of the ratio of flow setup cost to maintenance cost (F).

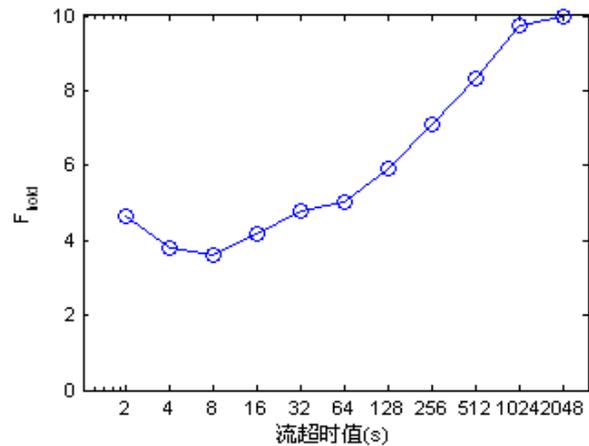


图 4.  $F_{hold}$  与流超时值的对应关系.

Fig. 4  $F_{hold}$  versus timeout value .

## 4 结论和展望

本文研究结果表明, 16 秒在新流和活跃流的量上都不大, 而且在一定的路由器处理任务和内存耗费的比例下, 16 秒是一个很好的平衡点; 在流平均重建次数方面, 它的值为 2.56, 首次下降到 3 次以下, 这表明现有网络设备的 CPU 已能够应付其重建任务; 16 秒的  $F_{hold}$  值为 4.16, 排名第三, 这说明了它在内存消耗方面也具有优势。因此, 本文认为使用 16 秒作为目前主干网数据流超时是合适的。

笔者认为, 今后的流超时研究, 除了需要从更多的角度寻找评估标准外, 另一个重要的任务就是为这些散落的标准建立一个综合评估数学模型, 以满足一般性的流研究; 同时也需要为一些特定的研究和应用创建专用的综合模型, 从而为它们提供更准确的流超时值。

## 参考文献:

- [1] CLAFFY K C. Internet Traffic Characterization. Dissertation for the degree Doctor of Philosophy[D]. University of California, San Diego. 1994.
- [2] CLAFFY K C, BRAUN H W, POLYZOS G C. A parameterizable methodology for internet traffic flow profiling[J]. IEEE Journal on Selected Areas in Communications, 1995, 12(8): 1481-1494.
- [3] RYU B, CHENEY D, BRAUN H W. Internet flow characterization: adaptive timeout strategy and statistical modeling[A]. Workshop on Passive and Active Measurement (PAM)[C]. 2001. 95-105.
- [4] SHAIKH A, REXFORD J, SHIN K G. Load-sensitive routing of long-lived IP flows[A]. Proceedings of SIGCOMM[C]. 1999. 215-266.
- [5] 周明中, 龚俭, 丁伟. 网络流超时策略研究[J]. 通信学报, 2005, 26(4): 88-93.

# Study on TCP Flow Timeout

WANG Yuan, DING Wei, GONG Jian

(CERNET Eastern China(North) Regional Network Center,  
School of Computer Science and Engineering, Southeast University,  
Nanjing, Jiangsu, 210096, China)

**Abstract:** Since drawn by K.C.Claffy in 1994, it has been well-accepted so far that a 64 seconds flow timeout value was appropriate for backbone network traffic measurement. As the Internet has been changed a lot during this period, is this conclusion still suitable? Based on a one-hour packet trace collected from a CERNET 2.5G province net border link recently, we did several statistics and analysis in the similar methodology as K.C.Claffy's. The conclusion was that a 16 seconds flow timeout value was large enough to provide an ideal effect of measurement. This conclusion would offer a reference for all kinds of engineering and studies related to network flow. In the aspect of experiment data volume, compared with K.C.Claffy's 1118 packets per second (pps), the volume of our trace was almost 650k pps, about 581 times of the former one. This paper would introduce the experiment and the course of engendering conclusion in detail.

**Key words:** network flow; timeout; backbone; network measurement