

组织: 中国互动出版网 (<http://www.china-pub.com/>)

RFC 文档中文翻译计划 (<http://www.china-pub.com/compters/emook/aboutemook.htm>)

E-mail: [ouyang@china-pub.com](mailto:ouyang@china-pub.com)

译者: lou\_goodman (lou\_goodman lou\_oxygen@163.net)

译文发布时间: 2001-4-27

版权: 本中文翻译文档版权归中国互动出版网所有。可以用于非商业用途自由转载, 但必须保留本文档的翻译及版权信息。

RFC 872            September 1982  
M82-48

# 局域网上的 TCP 协议

(RFC872 TCP-ON-A-LAN)

M. A. PADLIPSKY  
THE MITRE CORPORATION  
Bedford, Massachusetts

## 一: 摘要

在这个文档里, 我们证明那些认为符合 DoD(美国国防部)标准的 TCP 和 IP 协议不适合局域网的观点是不正确的. 这个文档是和 M82-47, M82-49, M82-50, and M82-51 紧密联系的.

## 二: 主旨

这篇论文的主旨是否定那种认为 TCP 协议用在局域网是一种 wozzle 的观点. 为了达到这一点, 我们需要知道什么是 wozzle, 什么是 LAN, 什么是 TCP.

### 1: Wozzles(一种可怕的动物)

第一个问题相当简单[1]:

一个晴朗的冬日, Piglet 正在扫它家门前的雪. 他突然发现 Pooh 在不停的绕圈圈, 好像想着什么问题. 于是 Piglet 对他说:

“喂, 你在干什么?”

“猎取.”

“猎取什么啊?”

“跟踪难题.” Pooh 神秘的说.

“什么难题?”

“这正是我在想的问题,我在跟踪什么问题?”

“你觉得有答案了吗?”

“还没有.看那.”Pooh 指着它前面的地面说:“那是什么?”

“雪痕,动物的爪痕”Piglet 兴奋的尖叫到,“你不觉得这是一个 Woozle 吗!”

于是他们互相是对方相信这就是一个 Woozle,一种可怕的动物.他们害怕起来,直到 Christopher Robin 来告诉他们他们看到的只不过是自己的脚印.

其实,这就像我们害怕 TCP 协议用在局域网中一样.我们误解了协议和环境本身,除非技术上不允许.

## 2: 局域网

第二个问题就不那么简单了.一般来说,局域网是一种短距离(几千米),高传输速率(大于几百 kbps),低误码率的通信机制(子网).它首先使得计算机(主机)之间能够通信,;其次,允许网络打印机和 TCP 协议终端与主机通信,尽管这一点不是必须的.

这样,从原则上讲,主机是异类的.它们不是相同操作系统的叠加.为了达到 ARPANET 称作“资源共享”和 ISO 称作“开放系统互连”的目标,主机之间通过层次协议进行通信.通信方式是主机到主机(点到点)或广播方式.(在一些特别环境中,例如 Ethernet,吸引人的方式是广播;在另一些环境,例如 ARPA-和 ISO-标准的“Internet”,广播方式是如此的昂贵一直与它不可能被实现.)

对于局域网的实现介质和拓扑结构,我们没作任何假定.局域网介质可能是双绞线,CATV,同轴电缆,光纤等.如果介质采用处理器到处理器的总线,那这个系统更像分布式多操作系统.因微处理器不能用 ARPANET 或 ISO 的层次协议.像“PDSC”(the Pacific Data Service Center)和“NMIC”(the National Military Intelligence Center)系统都不是局域网.

局域网的拓扑结构可能是总线,环网,或星型.从这一点来说数字 PBX 是一个局域网,因为它有传输介质能提供资源共享和开放互连,虽然不能保证传输率和误码率.从拓扑结构来说,它是中心星型结构.

对于我们来说,局域网的诱人性质是它的高数据传输率和低误码率.显见,实现这个性质的传输介质不能用为长距离通信网络而设计的 TCP 协议(我们并没有把带宽的浪费归结为数据包头.[2],pp.1509f,提供了对传统的通信方式的反驳.)如果你想做的只是让一些终端连接上一些主机,你根本不需要任何网络协议,你所做的只能被称为 LCN,而不是我们讨论的 LAN.

## 3: TCP

第三个我们需要知道的既直接又微妙.完全依靠我们对 ARPANET 标准协议的理解:直观上来说,Figure 1 和 Figure 2 都需要表述清楚,它们意味着 ARPANET 标准协议并非一个整体,而是层次结构.为了更清楚的说明这个问题,我们认为:TCP 协议是一个主机对主机协议(近似等价于 ISO 标准协议的第五层).它的最显著的性质是提供了可靠的逻辑上的连接.(这一点稍

候又会提到。) TCP 协议的另一个显著性质是它是为 Catenet(或者说是 Internet)而设计的,也就是说连接在不同通信子网上的主机之间可以像 Catenet 上的主机一样通信。TCP 协议的其它性质,像数据流控制,逻辑连接管理等很容易设计啦。

因为 TCP 协议的独特的地址机制(就是说,它用二维整体来表示外部主机地址,因为它不知道主机是否和它在同一个子网上),它需要一个接口(IP 协议)来处理不同子网的信息。这种接口就是 IP 协议。尽管 IP 协议被设计为和 TCP 协议独立,但它为主机间的数据传输提供了基础。

为了处理不同子网的问题,IP 协议拥有下面的性质:IP 协议察看一张保存有离一个主机最近的子网的信息(如优先级,服务等级,安全标记)的表,从而决定一个因特网地址是不是在这个子网上;如果是,它就把信息发送给这个子网,否则,它把信息发送给网关(它有另一个 IP 协议模块)。可见,IP 协议处理因特网路由,而 TCP 协议处理因特网寻址。因为一些子网传输能力差,只能传输小数据包,IP 协议还应该能把大数据包分割成适合相应子网的大小。最后,IP 协议应该提供一种机制,使得它能被其它协议辨认出来。(这是靠层次原则来实现的:你不需要了解 IP 协议所传输的数据,只需要了解 IP 协议要干什么。)

现在看起来有一点麻烦,因为有太多的机制。(更完整的讨论,请参考文献[4])。但这些已足够证明我们的观点。一种未公开的协议是 UDP 协议-----用户数据报协议。UDP 协议更强调速率而不是正确率,它是不可靠的。对于 UDP 协议来说,任何一个数据包都建立一个逻辑连接。这样,如果你想多路传输数据,你应该用 UDP 协议而非 TCP 协议。

### 三: 局域网上的 TCP 协议

不管你的主机是否属于一个局域网,也不管你是否了解 TCP/IP 协议,如果你要与因特网上的另一台主机通信,你必须用 TCP 协议。如果你想做一些网络上的应用(ISO 标准协议上的第 5 层的部分服务和全部第 6,7 层),就需要 TCP/IP 协议,因为它通过逻辑连接提供了可靠性,数据流控制,次序控制等内容。但如果你的应用不需要 TCP 协议的性质,就不要用它,不管你是在干什么。如果你想跟同一个网络上的主机通信,你可以自己设计一个协议,但这通常是极其糟糕的。如果你想让自己的网络保持自然状态,TCP/IP 协议根本不会阻碍你。但应该提醒你,你的应用多需要主机对主机的协议,因此,除非你想搞一个真正并行的东西..... 否则,主机间的协议是必不可少的。

现在我要讨论性能问题啦。这是一个非常细节化的问题。应该指出一点:在考虑可靠性的前提下,许多人(包括我)都为 TCP 协议而汗颜。因为研究机构在好多种类型的主机上的实验表明:虽然 TCP 协议友好的容错性,但它只有 12% 的效率。这让我们考虑又没有必要换一种主机对主机间的协议来替代 TCP 协议。(如果有这样一种协议,它也应该有 TCP 协议的可靠性这个优点。因为你是做 TCP 协议,而不是 LCN,像前边提到一样。)

抓住这只 Wozzle!

### 四: 其它性质

下面将阐述 TCP/IP 的其它性质:

1. TCP/IP 可以用在两个完全不同的操作系统之间;
2. TCP/IP 已经应用好几年了;
3. IP 层不限制子网提供的接口协议(尽管其中的一些是一种浪费);
4. IP 层不限制它的使用者,只要子网提供路由(不像 X.25);
5. IP 协议的网关同样拥有 3 和 4 的性质;
6. TCP/IP 符合 DoD 标准;
7. 应用协议和文件传输协议(包括 Email)已经存在并被用在许多不同的操作系统上;
8. TCP/IP 是受到美国国防部协议标准的支持的;
9. 研究机构报告的最新数据是:

	SUBNET	SUBNET	PHONE LINE
实际值(kb/s)	300	1.2k	9.2
理论值(kb/s)	800	9.2k	9.6

因为性质 8,没有其它协议比 TCP/IP 更适合.特别注意上面的性质不该被误解为局域网上 TCP/IP 协议的性能分析的一种限制.(但你应该知道局域网中性能分析的困难:1. 用一种备选协议来测量比较困难;2.大部分主机不能产生如你所设想的高数据产生率.)这样就像是限制了 TCP/IP 协议的运用.

## 五: 其它有用的数据

到这里,我们已经把问题讨论清楚,但进一步分析主机的性能是合理的,如果仅仅想搞清楚前面的讨论.主机终端是刷新模式,每一屏的内容有 16ms.那么中断 1 秒钟能刷新多少次?我们知道,硬盘寻道时间是 17ms,数据从硬盘提取出来要 2ms,如果数据传输到网络有 1ms,那么加起来就要 20ms,尽管主机什么也没做.如果本地硬盘的 I/O 速率是 16kb,那么 1 秒钟,主机只能提供 50 屏/秒的刷新率,大约 800kb/s,这个速率符合 TCP 协议的速率(见性质 9).在一个实际的系统中,主机看起来达不到这个速率.(主机数据传输速率的分析是困难的,因为它依靠系统内部的性能.因为经典操作系统的本质,如进程间共享栈的需要,输入数据占据内存的需要时的不可能得到比你所能产生的传输率更高的数据传输率.)

## 六: 结论:

认为 TCP 协议不能用于局域网的看法是没有任何根据的.TCP 协议完全能用于局域网中,实现主机之间的通信.

## 七: 参考文献

- 
- [1] Milne, A. A., "Winnie-the-Pooh", various publishers.
- [2] The LAN description is based on Clark, D. D. et al., "An Introduction to Local Area Networks," IEEE Proc., V. 66, N. 11, November 1978, pp. 1497-1517, several year's worth of conversations with Dr. Clark, and the author's observations of both the open literature and the Oral Tradition (which were sufficiently well-thought of to have prompted The MITRE Corporation/NBS/NSA Local Nets "Brain Picking Panel" to have solicited his testimony during the year he was in FACC's employ.\*)
- [3] The TCP/IP descriptions are based on Postel, J. B., "Internet Protocol Specification," and "Transmission Control Specification" in DARPA Internet Program Protocol Specifications, USC Information Sciences Institute, September, 1981, and on more than 10 years' worth of conversations with Dr. Postel, Dr. Clark (now the DARPA "Internet Architect") and Dr. Vinton G. Cerf (co-originator of TCP), and on numerous discussions with several other members of the TCP/IP design team, on having edited the referenced documents for the PSTP, and, for that matter, on having been one of the developers of the ARPANET "Reference Model."
- [4] Padlipsky, M. A., "A Perspective on the ARPANET Reference Model", M82-47, The MITRE Corporation, September 1982; also available in Proc. INFOCOM '83.

---

\* In all honesty, as far as I know I started the rumor that TCP might be overkill for a LAN at that meeting. At the next TCP design meeting, however, they separated IP out from TCP, and everything's been alright for about three years now--except for getting the rumor killed. (I'd worry about Woozles turning into roosting chickens if it weren't for the facts that: 1. People tend to ignore their local guru; 2. I was trying to encourage the IP separation; and 3. All I ever wanted was some empirical data.)

NOTE: FIGURE 1. ARM in the Abstract, and FIGURE 2. ARMS, Somewhat Particularized, may be obtained by writing to: Mike

Padlipsky, MITRE Corporation, P.O. Box 208, Bedford,  
Massachusetts, 01730, or sending computer mail to  
Padlipsky@USC-ISIA.