

# 一种新的网络传输协议—XTP

贾亚军 (新疆石油地球物理所 830011)

**摘要:** 本文是编译者在对 Alfred C Weaver 先生《The Xpress Transfer Protocol》的初步学习和认识基础上整理的。该文对 XTP 发展背景、功能及其前景等做了较详细的论述。编译者在网络的管理和维护中所遇到的诸多的协议问题, 在 Weaver 先生的这篇论文中得到了答案及思路。

**关键词:** XTP 协议 分布计算 分析应用

传统的 TCP 和 ISO TP4 已渐渐地不能适应人们对信息技术的高要求。另外, 各种各样的协议之间的转换犹如人类中众多的语种一样, 需要翻译才能得以沟通。这对信息产业的分布计算、分布应用等形成了无形的障碍。XTP (Xpress Transfer Protocol) 意在弥补这些不足, 以适应现代信息社会的需求。

## 一、XTP 的由来

一个国际研究专家组专门为现代分布计算系统设计、开发了一个新的传输层通信协议, 该协议是 OSI 参考模式中的传输层 (Transport layer) 和网络层 (network layer) 的集成。XTP 表示一种由若干种有突出特点协议发展而来的最佳观念的综合, 然后以新的服务将这种观念加以扩充, 目的在于支持分布式应用和嵌入式实时系统。因此, XTP 可提供独特的服务组合, 包括快速连接设置、大数据流水线能力、可选的应答、可选的再传输、流、纠错、速率和突发控制、消息优先级、协议内的调度、可靠和非可靠路由、外频带数据传送、多询址格式 (包括 Internet 和 ISO 8348 网络层地址) 和可靠的传输多端点 (Multicast) 等。XTP 的设计很明显地把策略与机制分离开来 - XTP 提供多通信机制, XTP 用户可选择那些支持用户期望的通信的机制。

XTP 的新服务为九十年代所期望的应用类型显示出其特有的价值: 计算机与高分辨率显示器之间的通信; 远程医疗成像; 用于事务处理的客户机/服务器系统; 高达 Tb 的数据库传输; 用于航海、航空、太空运输及工厂的实时控制系统以及多媒体电话会议等。举一个 XTP 应用的实例, 美国国防部已选择 XTP 为其 SAFENET (Survivable Adaptable Fiber Optic Embedded Network) 的协议之一。SAFENET 是 DOD 军事战事临界通信系统的下一代规范。

在图 1 中, SAFENET 需要两组协议。用于非实时系统的 ISO 协议提供传统的服务, 诸如文件传输、电子邮件和目录服务等; “Lightweight” 服务是在 XTP 上运行的。这些服务支持

应用于军事实时系统中的高吞吐量、低延迟及优先级控制。

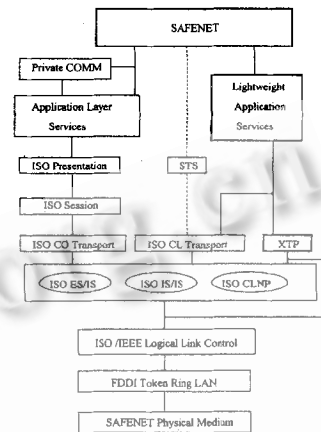


图 1 SAFENET

## 二、XTP 的促进因素

计算和控制系统的特征正在改变。对单一处理器的限制已不存在, 应用分布在多计算机环境中。因此在总体的系统结构上, 通信网络成为关键因素。对现代的应用而言, 通信协议的选择关系到服务用户能提供的服务功能及最高性能。

任何新的分布式系统协议都应具有这样的优势, 即比现有的协议提供更多和更好的服务。例如, 典型的传输协议提供或者完整的可靠服务 (面向连接), 或者非应答式服务 (非面向连接), 在二者中间不能有其他服务。一个较好的策略是提供各种机制, 使用户可选择哪种通信方式最适合其特定的应用。XTP 可做到对事务、用户定义的服务可靠性、多点通信和延迟控制的支持。

在性能方面, 任何新协议应允许各种各样的要求 - 从所有的软件到所有的硬件。这样系统设计者可在性能和经济方面平衡。在协议本身必须保持恒定时, 它的应用可从 Mb 大幅度地扩大到 Gb 规模, 扩展到 LAN、MAN 乃至 WAN, 而不需要替换。

## 三、变化的原因

网络界已推出了很多的传输协议, 著名的有传输控制协议 TCP 和 ISO 第 4 类传输协议 TP4。这些协议在当今都在使用, 为什么要改变?

改变的促进动力是由于网络世界在过去的十年里发生了

翻天覆地的变化。TCP 和 TP4 允许介质位错速率在  $10^{-5}$ , 现代光纤网络可提供  $10^{-12}$  或更小的位错速率; 很明显这会影响到再传输策略。同样, 早期的协议适合于低速率传输(即在专用电话线传输 56Kb/s)。当传输速率从 Ethernet 的 10Mb/s 提高到 FDDI 的 100Mb/s 甚至到 ATM 的 Gb/s, 点到点的位流比以前更密集。在这样高的速度下, 具有错误控制和错误免除的新机制则变得非常重要。

#### 四、功能优势

XTP 为现代的分布式系统提供了很多的功能。它可容纳大数据流水线、通过速率和突发控制防止缓冲溢出、管理大量的消息优先级结构、传输外频带数据、有可选的错误控制选项、允许选择流控制选项、对协议机制分离通信策略、避免网络层碎化以及给传输层提供可靠的多端点能力。以下将叙述这些功能。

##### 1. 数据流水线能力

随着技术的发展, 在通信应用中的分布式系统将会在数据流水线上传输大量的数据。当网络技术从 Ethernet 发展到 FDDI 和 ATM, 位流密度的增大是不可避免的。XTP 面对传输速度和数据的日益增长, 支持一个 32 位的顺序空间, 这样可允许一个 4GB 的滑动窗口。XTP 也可利用内部的同步域来支持其顺序数域, 从而达到 64 位的顺序数; 这使对未来的 Tb 网络的集成成为可能。

##### 2. 防止错误

传输协议采用 CRC 使应用避免介质错误, 用传输检查和 (Transport Checksum) 避免协议、节点或路由错误。XTP 不仅可提供这些服务, 还意识到随着光纤技术的应用及其低位错速率的特点, 由于缓冲区上溢出和系统内部拥挤造成的包丢失的可能性要远远高于由于介质错误造成的包丢失。因此, XTP 采取了速率控制 (Rate Control) 和突发控制 (Burst Control)。当末端系统处理其连接的服务质量 (QoS) 时, 接受者通过施加最大速率 (B/s) 和最大突发 (B/突发) 给发送者来调节传输关系。这些参数放在跟随有数据的路径中, 网络路由器动态地调整这些参数, 以防止拥挤。

##### 3. 消息优先级

多数数据链接层协议提供优先级机制, 这种机制可用于鉴别用户数据重要性的类型 (Ethernet 是一个典型的例外: 它不具备这种能力)。XTP 把这种鉴别很明显地放在用户消息级。每个消息可以 32 位标明其重要性, 发出该消息的网段所传输的所有包均具有该消息的优先级。在每次调度时, 节点对其最重要的包进行选择和操作。因此可以说, 考虑到一个包的传输时间粒度, XTP 的执行总是致力于其最重要的包。另外, 这种操作限定始终如一, 即在所有的内部路由器中都是可操作的。

这对高延迟控制要求的实时应用是必要的。

##### 4. 外频带数据

有时, 有关数据流的信息的发送不是以将其嵌入数据流的方式进行是很有益的。作为一个选项, XTP 可使每个包载有 8 字节的标记数据 (Tagged Data)。标记数据从发送端被送到接收端, 它向接收端表明它的存在, 但 XTP 从不对它做解释。标记数据对高层协议 (如: 表示层 Presentation Layer) 很有用, 它可提供有关数据的格式信息。标记数据还提供便利的方法来表示已被传输数据的计时。

##### 5. 可选的错误控制

XTP 提供了三种错误控制类型。传统的错误控制 (传输层检查和) 可保障正常的数据和可靠的、有检错功能的传输。作为一个选项, 接收端可生成一个快速的被动应答来加快再传输。这对于那些把顺序外数据认为是被丢失的数据的网络环境 (如 LAN) 是非常有用的。另一个选项是无错方式 (Noerror Mode), 它暂停正常的再传输。将接收无误的数据进行正确的排序, 但不再传输数据的间隙。这对数字化的音频和视频是有实用意义的。

##### 6. 可选的流控制

它提供三种具有错误控制的、正交的选项。传统的流控制适用于一般数据。保留方式 (Reservation Mode) 运行恒定的流控制策略, 接收端仅对这次连接给缓冲区发出应答; 这确保数据不因目的地的缓冲不足而丢失。第三个机制是使用无流 (No-flow) 选项来完全地禁止流控制。应注意的是, “自由流 (Free Flow)” 或 “流 (Streaming)” 操作方式在其他的协议中是不适用的, 这种方式对多媒体应用是很有益的。

##### 7. 策略和机制

XTP 经过很长的路途把策略从机制中分离出来。XTP 假定用户知道如何更好地优化连接参数。换句话说: 协议设计者至少了解任何一个单独应用是如何使用协议的。所以设计者的目的是提供能力而不是施加策略。这种分离的例子如上述的 “可选的错误控制” 和 “可选的流控制”, 后叙的 “可选的应答” 和 “可选的再传输” 也是这种实例。

##### 8. 最大的传输单元 (MTU) 检测

除了突发 (Burst) 和速率 (Rate), 发送端、接收段和内部路由器都可调整最大的传输单元 (MTU) 容量。经调整的 MTU 是路由上所有 MTU 的一个最小值。例如 Ethernet 的 MTU 为 1500B, FDDI 为 4500B。因此, 具有 Ethernet 和 FDDI 网段的路由的 MTU 为 1500B。了解了路由的 MTU 并使用该 MTU 作为传输层的最大的包容量, XTP 可避免整个网络层的碎化。

##### 9. 多端点 (Multicast)

传统的传输协议做了这样的假定, 即在两点之间建立双向的通信。但在控制系统变成分布式, 对组通信的需求变得尖锐。XTP 的多端点允许任何一个发送端向任意的一个大接收

端组发送可靠的数据。在多端点范围内,“可靠的”传输意味着接收端组的每个活化成员将接收到数据,任何数据错误都将由协议透明地恢复。尽管 XTP 的定义有意不包含多端点组管理,类似动态合并、脱离、再合并一个正在服务的多端点通信的操作可分别地进行处理。多端点对同时修改多个客户是很理想的,对诸如感应数据分布或维护多个一致性的数据库拷贝等应用也具有明显的实用意义。可靠的传输多端点在 XTP 是最具创新的特点,很有讨论意义。

## 五、性能优势

除了功能的增强外,XTP 还具有较高的性能,尤其发挥在网络接口集成于处理器的场合。通过对包头(Header)/包尾(Trailer)的设计、包头选件位的定位、连接设置及地址解释的优化、传输层及网络层融为一个传送层体系结构的集成、可选的再传输和可选的应答的利用,以及在长字节线上的数据装配等使性能得以增强。最终使 XTP 的执行能力提高。

### 1. 包头(Header)/包尾(Trailer)协议

常规的协议是包头协议,意指所有的协议控制和检错信息都在包头中。这意味着在传输前,传输检查和必须完成并插入到包头中。并且在启动传输前,消息必须先完成。XTP 将 4B 检查和放在包尾,故允许检查和在传输过程中由相应的硬件完成计算和附加等操作。把检查和移入包尾可使数据“流(Streaming)”贯穿协议,并消除了一次数据拷贝操作。

### 2. 定位的控制信息

所有的协议都允许用户在一组选件中作出选择;但这使包头的控制信息具有可变长度。通过对硬件执行分析,XTP 在包头定义选件位。不论选件是否被选,选件位的位置是固定的。这个定义简化了协议在硬件或软件方面的分析。

### 3. 连接管理

在 TP4 中,可靠的数据传输需要三次握手(六个包):两个用于请求和证实连接,两个用于发送和应答数据,两个用于释放连接和证实断开。图 2 表示 XTP 是如何用三个包交换完成打开连接、发送数据、应答数据及释放连接等过程。另外,XTP 允许传输的第一个包运载数据。

与传统协议不同,这种手段提前了成功的连接设置,使数据立即得以处理。

### 4. 地址解释

传统的网络层协议发送每个包时都带一个全网络地址,因而降低了整个网络的效率。通常的弥补措施包括高速缓存方法或包头前置量算法。当同时使用的网络地址数量比较大时,第一个措施变得很昂贵;在网络交通是随机性时(如 X-window 服务器),第二个措施难以提高网络性能。相反,XTP 利用一个基于键(Key-based)的方法,将全网络地址转化成为一个

32 位键(Key)放在建立连接的第一个包中。当第一个包处理后,所有其他的顺序包仅运载该键,并将该键作为查询表的简化索引。

### 5. 网络寻址方式

XTP 没有引入网络寻址方式。连接的第一个包含有一个参数化的网络地址,它包括一个地址类型和一个地址段。网络地址是通过参考被选的地址类型的寻址方式分析得出。目前已可辨别 12 种不同的寻址方式,包括 Internet 和 ISO 8348 网络地址。其他流行的类型可视需要增加。

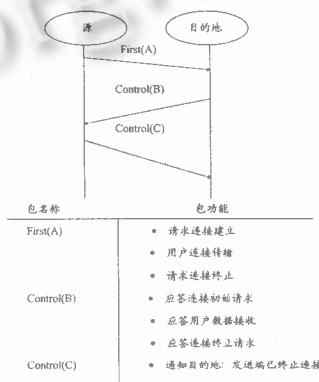


图 2 XTP 连接管理

### 6. 传送层(Transfer Layer)体系结构

XTP 是传送层体系结构,它仿造法国为实时 LAN 制定的军事标准,GAM-T-103。传送层体系结构衡量提供同层点对点可靠性和网络路由的效率。把传统的传输(Transport)和网络层合为一个传送层,在 XTP 是最有争议的设计决定,恐怕得不到标准化组织的支持。不论怎样,就目前而言,XTP 从这个结合中获得了高速度。

### 7. 可选的再传输

大多数再传输方式是 go-back-n。即若检测有错,传输窗口被复位,从被丢失或出错的包的第一个字节重新开始传输。因此,在一些情况下,go-back-n 可能再传输已被正确地接收的数据。XTP 可使用 go-back-n,或可选的再传输;对于后者,接收端应答已正确收到的数据的间隙,发送端仅再传输间隙。用户可视情况选择用哪种方式(再调用可选的再传输可使无错方式(Noerror Mode)完全失效)。

### 8. 可选的应答

每个包的应答是这样提出的:(1)假定网络经常丢失包;(2)发送端要求应答;(3)把机制嵌入策略。XTP 允许用户选择是否需要及什么时候需要应答。只要发送端请求应答,就能

得到应答。因此, XTP 允许用户从总是应答、有时应答和从不应答等范围选择。接收端的应答与数据到达或窗口的大小无关。只要发送端想知道连接的状态, 它就发问; 这时接收端响应, 将它所了解的现状通知发送端。

### 9. 数据合并

尽管数据合并是一个简单的概念, 但 XTP 从中获益非浅。首先, 包头/包尾的主要域被合并成 4 字节界线; 这使为恢复一个域需要的存储器存取次数变得最小。第二, 尽管传输的帧 (frame) 需要从 4 字节界线开始和结束, 但用户数据本身不需要从 4 字节界线开始和结束。偏移量 (Offset) 识别信息域中数据开始点, 长度域 (Length) 识别其结束点。当用户数据比最小帧还小时, 这种技术可发挥出很简便的处理作用。

## 六、XTP 的实际应用

XTP 若干个实际应用, 多在美国和欧洲。例如, 德国的 Braun 把 XTP 成功地用于多处理器体系结构。Henrichs 证实把 XTP 协议应用在并行化的可能, 柏林理工大学在 ISDN 上用 XTP 开发出功能很强的程序。在美国, XTP 已应用在多处理器的环境, Virginia 大学已公布了很多成功的应用。

### 1. 吞吐量及延迟

任何协议都期望在 CRAY 上显示出高性能, 但 CRAY 终究不是一个通用的平台。在以微处理器为核心的个人计算机和工程工作站上所具有的性能则更现实。在一台配有 50Mhz Intel 486、EISA 背板总线和双环 FDDI 的网络接口平台上所进行的实验表明可获得 38Mb/s 的点对点的传输速度 (即从用户存储器到用户存储器)。从用户存储器到用户存储器对 256B 的消息的传输延迟测试结果不到 1ms。

### 2. 多端点船舶航行数据

过去的研究项目为商船开发了实时通信网络, 它提供了所有的通信 (自动驾驶仪、电子罗盘、雷达、海难警告系统和船舶操纵)。Weaver 先生在实验室模拟波斯海湾上自动驾驶或人工驾驶的航行环境。这个环境是多端点的, 每次模拟都向多端点组发送航海信息 (航行路线和速度)。在演示中, 接收端 (放在船内的计算机) 可完成加盟到多端点组中、实时地显示船舶位置、脱离这个组再加入另一个组并很快地生成新船位显示。这种加盟、脱离、再加盟多端点组的能力显示出多端点传输的可靠性、高速度和多功能性。

### 3. 多端点文件传输

利用多端点文件传输, 一个单发送端可建立一个多端点组并以一次传输向所有的接收端发送。在系统中建立一个接收端组, 让发送端发送某一目录中的所有 ".c" 文件。所有的接收端都同时收到文件, 并且所有的包都透明地、正确地表明位错或缓冲溢出。接收端任何时候可加盟、脱离、再加盟多端点组, 不影响另一个组成员的传输可靠性。

### 4. 多端点 VGA 显示传输

对 XTP 的存储器到存储器测试是在一个 VGA 屏幕生成一个文本映象, 然后将该映象传输到多端点接收组。实验表明生成一个彩色的多边形映象, 把它传输给一个任意的大多端点组。在 FDDI 上用 50MHz 的 Intel486 平台传输, 多端点 VGA 屏幕传输速率高于 100 帧/s。

### 5. 多端点数字化 TV 映象

另一个存储器到存储器传输的演示是采用一组从普通 TV 摄像机摄取的、已数字化的静止帧。这组映象存放在磁盘中, 经解压缩后移入存储器。发送端 (FDDI 上的一个 50MHz 的 Intel486 平台) 发送连续的映象流, 所有的接收端捕捉到映象流并显示在它们的屏幕上。分布的传输速率约为 25 帧/s。

### 6. 数字化的声频分布

在一个有 FDDI 的 25Mhz Motorola 68020 平台上, 利用 XTP 编程把数据移入一个多通道数字化的声频分布系统中执行。实验表明用这套设备支持 40 - 110 个声频通道 (每个通道 64Kb/s), 这取决于一个 XTP 包的声频数据采样数。对一个 1024 声频采样的 XTP 包, 点对点的延迟为 5ms。很有趣的是, 多端点在所有测试中非常出色 - 一个多端点声频包的点对点延迟, 甚至一个接收端组仅有两个成员的延迟都比两个串接的独立端点的合并延迟要小。整个系统的数据容错能力非常突出。在声频数据包有 10% 被丢失的情况下, 由于 XTP 有效的错误恢复机制的作用, 平均的点对点的延迟增加不到 0.1ms。因此, XTP 可为传送层 (Transfer Layer) 数据分布系统以最小的整体延迟提供极具潜力的性能。

### 7. 多媒体

目前, 已开发出在 FDDI 网络上应用 XTP 的 DOS 设备驱动程序。利用该驱动程序, 可 mount 或 unmount 网络中任何磁盘, 把远程磁盘当作本地磁盘使用。采用多媒体接口, 可从任何输入设备 (摄像机、录像机或 LD) 记录视频和音频流, 并把合成的 A/V 流记录到网络上任何磁盘上。该系统可将全帧、全彩色、全速率视频 (30 帧/s) 和全同步音频提供给网络上任何 A/V 输出设备。测试结果表明, 通过 XTP 和 FDDI, 人们很难辨别出回放 (Playback) 文件是在本地磁盘还是在远程磁盘。

## 七、结束语

XTP 对传输协议设计是一个新的挑战。将当今分布式计算机系统所需的功能与现有的、颇具特点的协议加以组合, XTP 在计算应用领域从工程工作站到实时控制系统的应用前景是很广泛的。

那么 XTP 是否会被国际标准组织接受? 它是否会在网络通信和分布式处理系统取代传统的协议? 到那时, 我们已采用的依赖于传统协议的应用又会面临什么样的结局? 这些问题迫使我们不能拭目以待, 应积极地接受新的挑战, 紧紧跟上新技术的发展步伐。

(来稿时间: 1996 年 11 月)