

国家重点实验室高速局域网论证与实施

谢照华 (中科院科学和工程计算国家重点实验室)

一、安装高速网的必要性

科学与工程计算国家重点实验室承担着科学与工程计算领域中具有重大意义的基础理论研究,解决重大计算及科学可视化问题。网络分布式并行计算处理中必须在相关的任务间交换数据,进行信息传递、高速网络技术使工作站之间的进程通讯加快,提高并行计算程序执行的速度,加入网络计算的计算机数目的增加对网络速度及可靠性提出了更高的要求。计算机网络性能的飞跃发

展为采用网上计算机群并行计算来实现大规模计算问题及多媒体应用开辟了新的前景,保护了计算机与网络的投资。实验室于 95 年 9 月建成采用最先进的网络技术异步传输模式 (Asynchronous Transfer Mode - ATM) 将十八台工作站和服务器等组成的一个 ATM 异种机高速局域网与三个以太网互连的网络系统,参见图 1。ATM 与 Ethernet 互连局域网络系统。该系统的建成也为 863“可扩展并行机群系统”课题提供了研究环境。

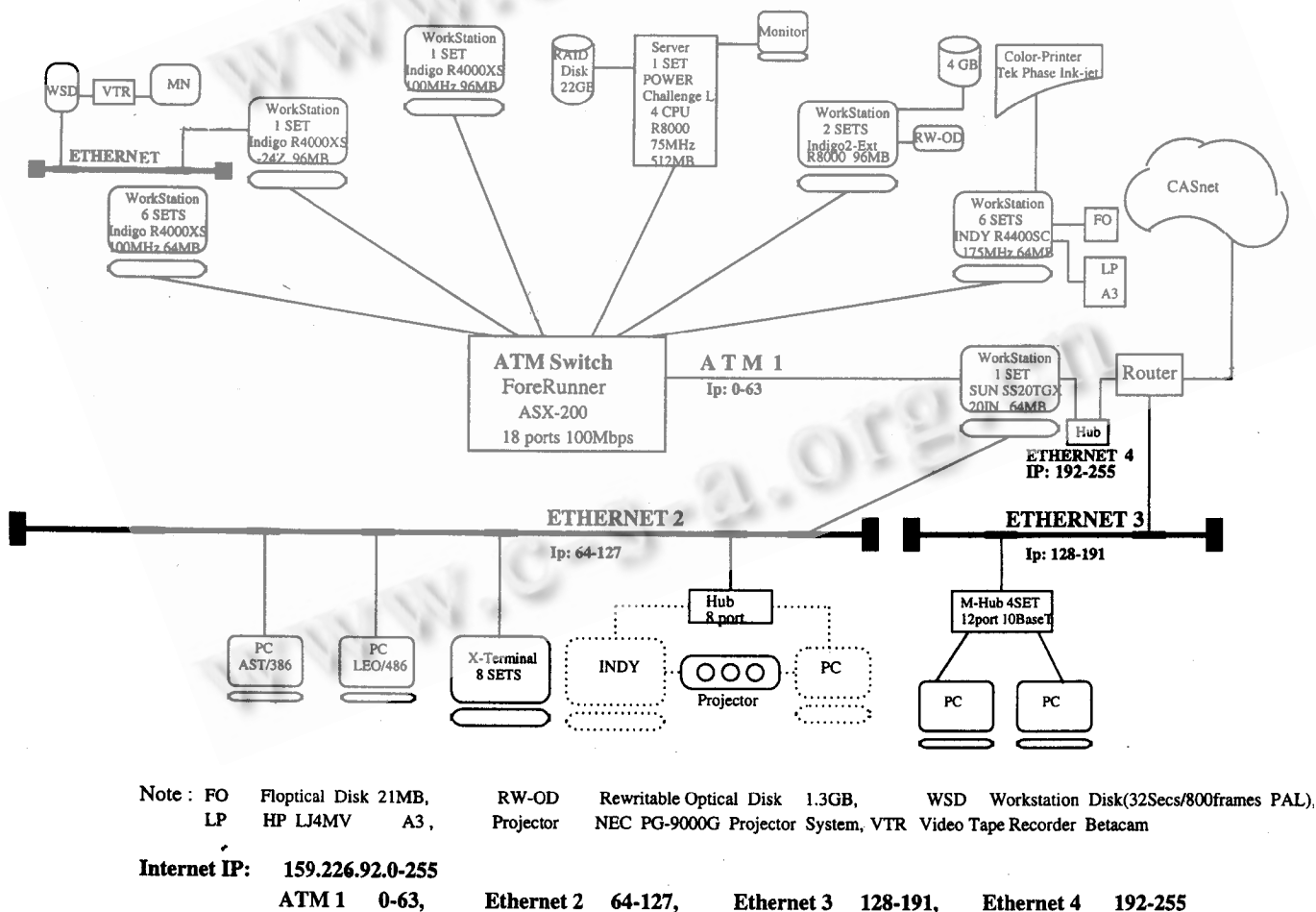


图 1 ATM 与 Ethernet 互连局域网络系统配置图

二、高速网络性能分析

1. 路由技术与交换技术

(1) 路由技术。当前的网络通常是由三个基本部件:智能集线器(Intelligent Hub)、网桥(Bridge)和路由器(Router)组成的。集线器是在局域网中提供一个可控制的存取点,允许多个以太网段存在于集线器中。网桥提供了高的数据吞吐率。而路由器则是建立与集线器互连的网络交换设备,网桥正被路由器替代。

在以太网和 FDDI 等这些基于包交换的网络中,路由器是用于局域网和广域网互连的最常用的设备。路由信息是包含在数据包中第三层网间网层的信息。第三层路由信息向每个包增加 40 个字节的路由信息,其中包含网络地址,以识别每个网络站点(用户)和它们在网络上的位置。网络节点地址必须正确配置,且与其连接的逻辑子网匹配。但路由器并不管每个网络站点,而只管该网络站点所在的逻辑子网。一个物理多协议路由器,实际上是由许多单独的逻辑路由器在起作用,每个逻辑路由器支持一种协议,有其路由表(路由拓扑数据库)。当有用用户连入网络或迁移,就需重新配置路由表,这种重新配置不但需要许多时间和高技能,且重新配置占用了很多的网络带宽。

路由技术提供了局域网和广域网的连接控制网段之间的通信量,以节省带宽。采用多协议路由器是局域网和广域网互联的有效方法,但路由器不具备可扩展性来支持大量使用带宽的应用程序及单用户高速专用局域网的连接,而面向连接的交换技术则比较适合。

(2) 交换技术。鉴于路由技术需要静态地在端点设备和路由设备中配置第三层路由信息,对大型网络系统其改变是困难的,同时新的网络应用如网络并行计算,客户/服务器及多媒体应用等要求提供足够的带宽来消除瓶颈,消除延迟和网络故障。故而采用交换技术支持计算机之间多条专有的高速连接,为大型网络提供强有力的可控性和可扩展性的虚拟网络正在兴起。虚拟网络是由软件控制和管理的交换机组成的动态网络,如 ATM 网络,交换以太网和交换令牌网等。

2. 各种高速局域网分析

(1) FDDI 光纤分布式数据接口。FDDI 是标准化的,广泛使用的高速网络技术,传输速率为 100Mbps,网络拓扑结构为环形,使用令牌传递协议。物理层需要标准定义的 FDDI 适配卡及电缆可采用第 5 类非屏蔽双绞

线(UTP-5),多模光纤(62.5/125 μm)或单模光纤。可以作为以太网和令牌环网的主干网,环形网周最大为 100 公里。中关村科学与教育实验网的主干网就采用了 FDDI 双环网,其容错性能很好。但 FDDI 技术存在以下不足之处:

①FDDI 是共享通路网络存取机制,虽然速度可达 100Mbps,但随着网上用户的增加效率也降低,而且 100Mbps 已是极限速度。

②FDDI 与以太网的桥接和路由选择的复杂性。FDDI 网中的包格式与以太网中使用的包格式完全不同,在这两种网间进行信息传送就必须在网桥或路由器中执行包格式转换,这样 FDDI 的网桥与路由器的价格很高。

③FDDI 使用的是可变长的包传送及共享通路传送,其不适于信息量大且要求实时的多媒体视频/音频的传送。

(2) ATM 异步传输模式。ATM 是建立在 ATM 交换机上的高速网络技术,可以根据网上计算机的需要来分配专用带宽和在网络站点安装相应的适配卡,局域网 ATM 常用的速度为 25.6,44.736,51.84,100,155 和 622Mbps,速度在 100Mbps 以下可采用无屏蔽双绞线,100 和 155Mbps 可使用多模光纤,622Mbps 则需使用单模光纤。

ATM 最初是由美国电话公司开发的,数据是以确定的长度 53 个字节数据包构成一个 ATM 信息单元(cell: 5 个字节头标,48 个字节数据)。采用信息单元便于实现质优价廉的每秒传输几千兆位的交换机。ATM 交换机与网上站点的连接,以及交换机之间的连接可以按多种速度操作,这些不同速度的链路之间速度的匹配是 ATM 交换机本身固有的功能。

重要的是 ATM 可以为面向连接的应用提供各种新的网络应用,其中包括电子邮件,客户/服务器,交互式的对等数据应用,以及声音和视频的传输。对于面向连接的传输,则需通过交换建立“虚连接”,即两网络站点间的路由。ATM 单元交换虚连接的建立是比较简单的,交换是基于头标信息中的虚拟电路识别标志(VCI)字节,现在每个数据包中都包含了全球统一的路由信息,这个通用的全球规范。

ATM 是一种灵活的技术,适用于从工作组局域网到广域网的各种应用,但这种虚拟局域网只在网络模型的第二层上工作,其不能解决第三层相关的问题,故 ATM

虚拟局域网与其它现有的局域网和广域网连接时,还需要路由器。同时使用局域网仿真使局域网和已有的局域网,如以太网可以互操作,即在不丢弃已有的应用和系统的情况下,无缝地过渡到使用 ATM。

ATM 网络技术也有不足之处,就是 ATM 把网络上的负担的相当一部分转移到网络站点,如工作站和管理系统上,网络站点必须把信息包格式化为单元以便传输,并且在接收时将接收的单元信息转换回信息包(称为分段和重组 SAR),其可以由装在网络站点上的 ATM 适配卡上的协处理器来执行,但目前多由主机 CPU 来完成,占用了一部分 CPU 的处理能力,所以使用较高的数据传输速率(100Mbps 以上)的 ATM 网,才更为经济。

(3) 虚拟局域网 (VLAN)。虚拟局域网是用基于分包交换机的网络替代基于路由器的一种新的网络基本结构。其容许以太网、令牌环网和 FDDI 网段享有与 ATM 环境中同样的交换性能。交换机基本上说是一些具有更多功能的多口的高速桥接器,所以交换技术就是将网络扩展开来。在进行信息交换时,虚拟局域网的功能是“重复生成”路由。VLAN 的一个明显的优点是具有从 VLANs 到在 ATM 中的仿真 LANs(ELAN)的转换功能,这样 VLAN 容易实现向 ATM 转移,而不需要进行人工处理配置 LAN 仿真配置服务器。

LAN 交换机提供建立 VLAN 网所需的硬件和软件,使在网上具有传送包的能力。这种结构的优点是容易改变网络,用户不需改动电缆线便可以逻辑地在工作组之间移动。也就是用户可以物理地在网上移动到任何地方但仍保留与网络的连接,只要有访问所有权。由于 VLAN 基于快速交换技术,因此不论从性能上还是资源的有效利用上都有提高。VLAN 可以逻辑地定义工作组,满足在动态工作组内增加带宽的同时,创建动态工作组以满足开发的需求。

这种把路由器和 ATM 单元交换的优点结合起来的新技术,可以较经济地满足数据速率在 100Mbps 以下用户对频宽的需要。在共享通路局域网若有 10 个用户,每个用户得到 10 的可用频宽。但在这种包交换局域网每个用户得到的是 100% 的各自局域网的频宽,网的总频宽容量将增加 10 倍。而且可以支持全双工功能,以支持对等应用服务。还有不需要改变原有网络站点系统、电缆及应用程序。LAN 是对 ATM 的一种技术补充,满足目前一些应用对容量、可靠性和安全性日益增长的要求,并逐渐转到 ATM 单元交换网络。

(4) 100Mbps 高速以太网。以太网传输速率为 10Mbps,信息传送可以复盖 2.5Km 的范围,有 1024 个站点,基本满足各种园区的需求。由于要提高信息传输速率到 100Mbps,而有 3Com 公司开发了“快速以太网 (Fast Ethernet)”和 HP 公司开发了“100VG - AnyLAN”网络,均可复盖几百米(现为 200 米)的范围,适合工作组的需求。

①100VG - AnyLAN。支持传输速率为 100Mbps 的以太网和令牌环网信息包格式,可以用 4 对第三类无屏蔽双绞线 (UTP - 3),两对屏蔽双绞线 (STP) 或多模光纤。传输协议是全新的,称为请求优先级存取方法 (Demand Priority Access Method - DPAM),有两个优先级“一般 (normal)”或“高优先级 (High - priority)”,高优先级的信息包是适合于要求时间延迟小的应用,如多媒体在 100Mbps 的带宽中享有更大的比例。今后会使用 2 对 UTP - 5 双绞线,全双工操作,随机连接和有更高的速率。其需要 100Mbps 的多端口中继器,和需要更换网络站点的网络适配卡。

②快速以太网 - 100Base T。支持用现有的以太网网络接口层,即第二层以 100Mbps 操作,而替代 10Mbps。可以用 4 对 UTP - 3 双绞线,两对 UTP - 5 双绞线,两对 STP 双绞线或多模光纤。今后会在两对 UTP - 3 双绞线上使用带有流控制的全双工操作,以获得较高速度。需要在网络站点上配置 100Mbps 的网络适配卡。

这两种以太网初次安装投资较少,特别是 100Base - - T 是现在 10Mbps 的最经济的升级方案。94 年第四季度 3Com 公司推出符合 100Base - - T 的网卡,95 年下半年才推出 100Mbps 集线器。总之 100Mbps 高速以太网是用来提高带宽的改革措施。

三. 高速局域网的建设

1. 高速局域网方案选择

实验室已有和准备安装的高性能图形工作站和服务器共十八台, X 终端八台, PC 机多台及如此多的其它硬件设备,要使这些资源充分发挥效益,更好地适用于分布式并行计算及科学可视化,就最好建造由这十八台计算机组成的工作组高速局域网。经过一段时间的研究与市场调查,基于本文前面提到的对各种网络技术的评价,认为采用美国 FORE Systems 公司的 ATM 局域网络技术是最好的解决方案。

ATM 网络技术被各大计算公司及整个计算机世界认为是本世纪末和二十一世纪工作站连网的最佳 LAN

解决方案,也是一种适用于各种网络基础结构的通用数据传输方案,可方便地与现有的 LAN 网连接,其又拥有一个通用的全球规范,ATM 对于数据、声音和从局域网到都市网及广域网通信的低延迟和按照需求提供带宽的功能,以及 ATM 做为一个虚拟网具有的可靠性、安全性和灵活性已被许多国家选做主干网。

ATM 具有如此多的优点和发展潜力,在 94 年 ATM 网络产品却只有美国和日本几家公司有可以商业化的产品,国际上 ATM 网络标准规范也未出台,各厂商生产的 ATM 交换机规格也就不相同,94 年 8 月美国 Dataquest 公司公布的 ATM 局域网市场占有情况看,美国 FORE Systems 公司占了市场的 64% 份额,但其 94 年第二季度财政收入只有 5 百多万美元。不过从第一台 ATM 交换机交付使用,这种技术何时变成现实就没什么可谈的了。何况各大公司都投入了大量的资金进行产品研制和产品陆续投入市场。我们对几家公司的 ATM 产品进行了调研,并与 FDDI 网络产品进行了性能和价格的比较。

(1) 要实现至少 16 个站点的数据传输速率在 100Mbps 以上的 ATM 局域网的 ATM 交换机,当时只有 FORE Systems 公司的 ForeRunner ASX200 可以实现,参见图 2 ASX200 无竞争时分交换构架,该交换机可连接传输速率为 100Mbps 的 24 个站点或 155Mbps 的 16 个站点。有的公司则需要 2 个 ATM 开关互连才能实现,这样价格高。

(2) ASX200 在由美国的 Data Communications 杂志和欧洲网络实验室组织的 ATM 测试中,其数据传输延迟时间最小,能很好地实现对时间敏感的大量数据在网上的传输。分析其结构是由于 ASX200 交换机中采用了时分交换构架(Time Division Switching),是一种无竞争(Contentionless)构架,在此种构架中,中央交换构架的传输能力大于或等于所有单个输入/输出端口的传输能力的总和,因此对端口一级来说是无竞争的(或称无阻塞 non-blocking)。

(3) 该公司也是唯一的一家提供 SGI 工作站适配卡、和不同版本操作系统支持的驱动软件的公司。在 ATM 产品还不太成熟的情况下,选择一家公司的 ATM 交换机和 ATM 站点适配卡会使其之间功能的匹配不成问题。

(4) 该公司能提供 LAN 仿真软件,以使得实验室不

能连接到 ATM 高速局域网上的设备连在 Ethernet 网上,实现从 Ethernet 局域网到 ATM 的无缝联接。

(5) FORE System 公司提供性能良好的 ForeThought 互联网络软件。并提供 ForeView 网络管理系统,这种基于图形的网络管理可以对 ATM 网络进行配置、控制、监督和查找故障。而且其与 SunNet 管理软件这种广泛使用的网络平台集成一起,可以使网络管理员对网络范围内的所有站点进行一体化的管理。

在价格上采用 ATM 网络比有的公司提供的 FDDI 的配置价格还低。94 年有报道 ATM 被认为是昂贵的技术,每个站点适配卡的价格为 2 - 4 千美元。但比较 Ethernet, FDDI 和 ATM 的 1Mbps 带宽的价格时,可看到 ATM 的 1Mbps 带宽约为 2 美元, FDDI 为 40 美元,而 Ethernet 则需 50 美元。从我们的实践中也感到 ATM 是实现高速交换最便宜的方法。

2. 建成 ATM 与 Ethernet 互连局域网络系统

将新采购的设备与原有的设备重新设计为 ATM 与 Ethernet 互连局域网络系统的配置考虑网上各种设备的功能、使用要求及网络通信负载,形成四个子网,其中 ATM 子网采用的是美国 FORE Systems 公司的局域网产品 ATM 交换机和工作站上的适配卡,将十八台工作站(包括一台 SUN SPARC 20 TGX)连接成 ATM 异种机网,传输速率为 100Mbps,在 SUN 工作站上安装了网管软件。并把 SUN 工作站做为路由器与实验室内的 Ethernet 网互连,和通过另一个路由器与计算数学所内的 Ethernet 网互连,以及联通到科学院网和国际上的 Internet 网。实验室内的 ATM 网是当前国内最大的 ATM 局域网络,为进行并行算法研究提供了一个非常好的环境,极大地提高实验室内设备的整体功能。同时对实验室内的设备重新做了布局,在每个研究室内安装有联网的 X 终端和 PC 机,研究人员有很好的工作环境。

实验室互联局域网自 95 年 9 月验收至今,系统运行稳定可靠。

参考文献:

- [1] Paulina Borsook, ATM peeks out, Interest to the people, it Data Communications, IEEE Spectrum, January 1994, 26 - 28.
- [2] Daniel Abensour & Barbara Leonared, The ATM Impact, it NetNews IBM, August 1994 3 - 5.
- [3] FORE Systems, Inc., Real ATM Network, Concurrent Technologies Corporation Report, May 1994.