

# 分布式处理机通讯结构

戴家林、徐炜民 等译  
蔡勤业、程 杰

白 英 彩 校

苏州电子计算机厂

一九八三年三月

# 译 序

小型机和微型机的迅速发展和各种通讯技术的日趋成熟，促进了分布式处理系统的进一步发展。分布式系统具有高计算能力、良好容错性能和分布处理等特点，近年来已引起了国内外普遍关注，近年来计算技术发展趋势表明：八十年代将是分布式处理系统蓬勃发展的十年。

本书从通讯结构的角度讨论了各种分布式系统的拓扑结构及其特点，研究了报文分组转接、线路转接、总线通讯等通讯方式，并介绍了典型实例。本书可作为高等学校有关专业研究生教材，亦可作为计算技术人员的参考书。

参加本书翻译工作的还有白英彩、范良根、王承惠、陆群英、周虹萍、杜轩华、潘学正、项湜伍和黄元庚等同志，全书译稿由白英彩同志校阅。在本书翻译过程中孔宪豪同志给予关切和帮助，在此表示感谢。

译 校 者

1983.3

# 目 录

|     |                        |        |
|-----|------------------------|--------|
| 第一章 | 分布式系统概论 .....          | ( 1 )  |
| 第二章 | 各种拓扑概念的通讯结构 .....      | ( 9 )  |
| 第三章 | 网络通讯、信件交换与报文分组转接 ..... | ( 26 ) |
| 第四章 | 互联网络和线路转换技术 .....      | ( 44 ) |
| 第五章 | 总线结构 .....             | ( 86 ) |
| 第六章 | 传输技术和数据链控制 .....       | (114)  |
| 第七章 | 系统举例及其概念 .....         | (128)  |
| 第八章 | 总结 .....               | (153)  |

# 第一章 分布式系统概论

## 分布式系统

分布式系统的定义很多,而且五花八门。在本书中,我们选用一种得到许多研究人员[1, 2, 3, 4, 5]赞同的简单定义。分布式系统是这样一种系统,在该系统中,处理负载(Processing load)系分布在若干处理单元(PE)上,这种单元定义为处理机——存贮器对,它们在物理上相互连接,但是允许相互进行数量不等的软件通讯。一般情况下,在单元之间传送信息会有一段时延。可以被认为是分布式系统的典型计算机系统有用多微处理机实现的通用计算机、具有能执行特定功能(如快速富里哀变换)的专用辅件的多计算机、以及ILLIAC IV一类的阵列机、多处理机和计算机网络。

分布式处理机领域的研究主要在三个方面取得了进展:网络系统[6],专用系统[7],以及高性能通用系统[8]。此外,这一领域的研究主要在两个方面取得了进展,即计算机通讯和计算机结构[6, 9, 10, 11]。本书是以结构的观点编写的。

分布式处理带来的主要益处有资源共享性,系统的高利用率、负载的均匀性,流通量的增加、响应时间的改善、较好的模式性以及容错性能的改善。我们并不是说能在同一系统中实现上述全部优点,但是可以说分布式处理所能带来的潜在优越性是巨大的[1, 3, 6, 8]。

在分布式处理系统的设计中,存在着三种主要结构性问题:(1)进程间相互通讯;(2)计算机间相互通讯;(3)应用进程和系统进程的划分与分配(application and system process Partitioning and assignment)[12]。进程间相互通讯是人们颇为熟悉的功能,因为它在Dijkstra和其它操作系统中有很强的背景[13, 14, 15, 16]。计算机间相互通讯也是一个人们相当熟悉课题。事实上,某些设计技术是服从分类图结构的[9, 10, 17, 18, 19]。通用分布式处理机的进程划分是最近才开始发展起来的,目前研究人员正在为这一课题构造词汇和术语。

本书对适用于分布式处理系统和网络的计算机间相互通讯技术作了详尽讨论。有关相互连接的设计,从系统结构设计到细微的数据链路设计,均包罗于本书之中。这个问题的重要性怎么强调也不算过份,因为即使研究人员是否努力将微处理机或小型计算机互连成分布式系统这类问题不在我们考虑之中,我们仍然会遇到总线设计,通道设计,I/O设计等等计算机间相互通讯问题。为了使本书更为实用,我们还讨论了计算机通讯所必需的通用结构技术。这样,小型计算机网络的设计人员就会象那些试图将主计算机相互连接起来的人员一样,认识到本书的实用性。此外,上述两种设计人员都会发现,只要仔细遵循文章中所叙述的各项技术,他们就能得到下述两种主要利益:

1. 本书为每一种设计问题均提供了若干种结构,所以如果一位设计人员考虑了本书所述的各种技术,他就可以信相,自己已经考虑了许多主要的设计方案。
2. 设计人员应该得到一个如何在结构上组织他的系统的总体概念(Perspective),因

为本书在结构上提出了一个用于网络和分布式处理机的计算机间相互通讯结构技术的概  
观。

图 1-1 给出了一个分布式系统分层观点的初步说明。由于分布式系统中有许多进程在  
运行，所以系统可以看成是通过一层硬件和软件规程(Protocol)，在子网络上进行通讯  
的一组协作进程。图 1-1 给出了一个分布式处理概念的简单图解，其中有若干处理单元  
(主机——主机处理单元)，它们经由通讯子网进行通讯。系统中有虚实两种通讯。一个特  
定的系统与图 1-1 所示的系统相比，可能拥有更多的层次。

在本书中我们要叙述通讯子网的结构，这在网络和分布式处理机的建造中特别有用。  
准确地说，本书有一章专门用来讨论在设计分布式处理机和网络时将会遇到的各种系统设  
计论点和模型(第 2 章)。还有三章从结构上详细讨论了分布式系统中通讯子网的三种最为  
常见的类型，就是信件和报文分组转接(message and packet switches)(第 3 章)，线路  
转接(第 4 章)和总线(第 5 章)。第 6 章讨论了实际硬件传输通路设计中的各种硬件细节。  
第 7 章讨论了一些系统范例，分布式处理系统方面未来的趋势则属于第 8 章讨论的内容。  
在本章的下半部分，我们将讨论数字通讯系统结构的演变，以及简单分布式系统的三个  
范例。

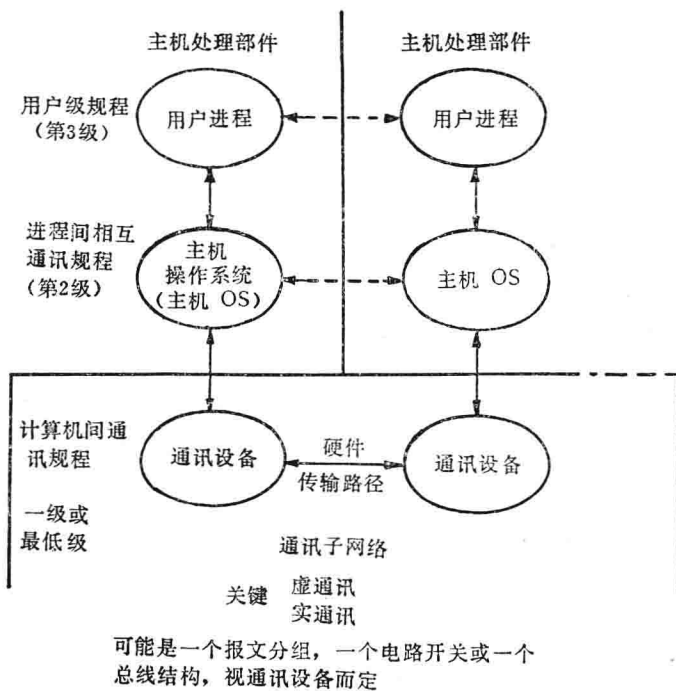


图 1-1 通用层次通讯系统模型

### 数字通讯系统结构的演变

在整个历史过程中，人类不得不进行通讯。人类使用了各种各样的通讯技术，从鼓信  
号烟信号一直到卫星数据传输装置。现代通讯技术是从电报到穿孔纸带电报装置到电传打  
字机一步一步发展而来的。

在过去的十五年到二十年中,随着计算机技术的发展,计算机系统也开始需要收发信息。由于分时系统和小型计算机网络的出现,计算机的用途增加了,通讯技术也更为重要了。由于某些数字总线概念十分复杂,代价也很高,因此各种更为经济(cost-effective)的通讯传输线技术便应运而生[3, 6, 8, 20, 21]。随着这种技术的发展,通讯技术也必须与计算机技术一起发展。通讯子网就是计算机和通讯系统之间的基本接口之一[1]。

最初人们认为计算机是中央设备,并严格地按照中央设备来生产计算机[17, 20]。所有外部设备都与计算机处于同一物理位置,即本机工作点。后来的发展(更为高级的外部设备组合体)使得某些外部设备能位于物理上的远程工作点。随着系统技术的进一步发展,小型计算机的网络允许不同的系统组件相互进行通讯[6]。

在计算机通讯的历史上,最初的步骤之一就是要将大量的外部设备连接在单个工作点上,如果操作系统具有适当的结构,我们就可以将各终端设计成具有相互作用的能力。只要我们愿意,各终端在物理上就可以是分布式的,这样,一个简单的分时系统就可以由一个具有批处理能力的本地工作点和若干远程终端工作点所构成。远程终端可以用总线或具有拨号能力的电话线同主计算机连接在一起。于是这种处理节点就可以被许多用户交互使用。每个用户均有分享系统资源的能力。可以进一步使远程工作点位于与实际计算机设备相距数百英里以外的地方。分时系统使得在学科上毫不相干的人们有勇气用计算机去解决形形色色的问题。

多处理机/多计算机系统是数字通讯技术发展的又一范例。典型的多处理机系统含有若干处理机,它们通过一个总线结构与一个共享存贮器进行通讯(图 1-2)。这种系统具有更高的资源利用率,并有失效保护失效弱化作用。它还提供了两种数字通讯技术:一种是从总线到存贮器的通讯,另一种是在存贮单元中存放信息供另一组程序以后读出。这种存贮器共享通讯概念在读写同步方面还存在一些问题,所以在计算机系统中要用到测试,置位和信号灯操作。同步问题在本书中不作讨论。在 Lorin[13]的著作中可以找到这一问题的详尽叙述和介决方法。图 1-3 中给出了一个类似的系统,它通称为多计算机,或联合(federated)系统。它的基本通讯手段是通过处理机之间的相互通讯链路。每台处理机均具有它自己的存贮器。

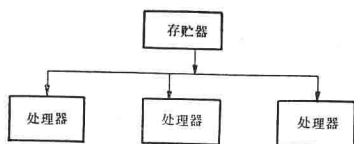


图 1-2 公共总线多处理机：分布式系统范例

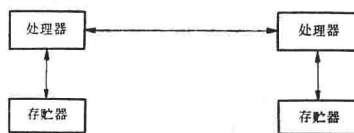


图 1-3 联合系统：分布式系统范例

对上述系统类型进行扩展,就成了图 1-4 所示的计算机网络或称资源共享网络[3,6]。我们可以将计算机网络描述成一个相互连接的计算机群。通常这些计算机至少能执行单台物理处理机所能够执行的两种不同功能。一种功能是主机所执行的通用处理功能,另一种功能是通讯功能和控制功能的合成。执行控制和通讯功能的单元系通称为信件转接或报文分组转接的存贮程序处理机[6]。

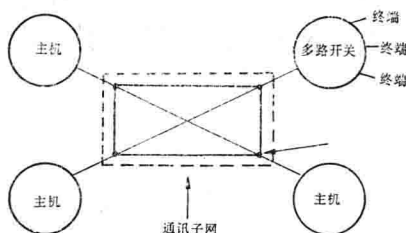


图 1-4 网络系统：分布式系统范例

网络的出现基于下述几种理由：

1. 有了计算机网络，用户就具有了对系统部分更新的能力，而不必每隔几年就用一台更大的中央计算机来取代原有的大型中央计算机。
2. 有了计算机网络，用户就可以要求本机工作点上所具有的各种各样的服务（例如 ARPA 网[9]中的 ILLIAC IV）。
3. 网络的总处理能力可以大于现有的任何一台计算机。
4. 网络的拓扑结构可以适合各种应用。
5. 有了计算机网络，就可以既方便又清楚地将用于复杂用途的软件拆成段，所以软件研制较容易。
6. 与单机相比，网络与它所具有的大量资源的容错性能更强。
7. 与巨型机相比，网络研制较容易。

并非所有的分布式系统都明显地具有图 1-1 所指出的全部规程级。最近，在实时军用系统中多微处理机的影响下，一些研究人员已开始注意专门化的分布式系统问题[1]。这些系统不同于网络，也不同于比较适用的分布式处理机，后者的通讯处理功能隐含地甚至明显地制成硬件，（在这种情况下，所有通讯道路均采用硬接线方式，所有通讯规程和通讯联络(traffic)均系事先安排好，并且是一成不变的）。这类分布式系统结构通常面向某种特定应用，它们与网络的不同之处在于：前者的应用可以是逻辑上的分布，而不是物理上的分布[2]。

网络的主要优点在于它们分享诸如物理设备，文件，程序，数据和外部设备等整个资源的能力，即使整个资源不在本机工作点也无关紧要。于是在本机工作点没有巨大硬件投资的情况下，也能提供巨大的服务能力。

在参考文献 1, 4, 11, 12 和 22 中，我们可以进一步看到从系统观点出发对分布式系统结构所作的全面论述。

## 通讯子网

所谓数字式通讯子网就是任何可用于互连数字装置的子系统，它可以简单到是一组计算机寄存器与计算机算逻部件之间的连线，也可以复杂到是一台用作为存贮——转发通讯开关的带有存贮器的计算机。在通讯子网设计中要作出决策至少要涉及以下几个主要方面：

1. 系统拓扑结构——整个系统的定义和结构，并决定究竟用通路（无需选路）还是用转接器（采用选路策略）来实现通讯子网。

2. 信件和报文分组转接设计——用通讯处理机还是用处理机系统来构成系统结构中的通讯结构, 执行信件转接或报文分组转接。

3. 线路转接设计用十字开关还是用分时多路开关来执行线路转接功能, 从而实现系统结构中的通讯结构。

4. 子网络总线或通路设计——用专用通讯传输线、共享总线, 还是用共享存储器来实现系统结构中的通讯结构。

典型的拓扑结构有环形, 星形以及点到点(point-to-point)连接系统。计算机的拓扑结构系用通路和转接器实现。通路是一种供信件传输的中间介质, 可以是专用数据传输线, 共享总线或公共存储器。然而, 由于总线的硬件复杂, 所以从设计的观点来看, 它们是最令人感兴趣的东西。通路设计中的主要决策要涉及到一台装置如何受通路控制, 以及信件如何在通路上传输等问题。转接器在信件的起点和终点之间进行干预, 以建立通讯路由, 它们可以是线路转接或信件和报文分组转接(系处理机或诸如 Bolt, Beranck 和 Newman(BBC)Plurilue IMP[23]一类的计算机互连而成的系统)。转接器最终必须将信息传输到如图 1-1 所示的硬件传输通路上。我们可以进一步把计算机间相互通讯层分为二层, 一层与转接算法有关, 另一层与数据通路控制有关, 这样我们就把图 1-1 分成了两部分。

在信口与信口之间的速度要求较高而选路决策简单的情况下宜采用线路转接。信件转接一般较线路转接为慢, 但允许有较复杂的选路决策。

## 小 结

本章简略回顾了数字通讯技术发展过程中的一些重要概念, 讨论了网络和分布式系统以及它们的应用和优点。在以后几章中, 我们将从通讯结构的观点来叙述网络的类型, 并讨论这些通讯子网的构成。

应该指出, 本书中所选择的分布式处理机模型系规定与设备(主机)无关。这样, 研究人员就能运用本书所述的技术研制若干种形式的小型计算机网络, 或应用这些概念改进多微处理机系统的通讯方案。研究人员的责任是在所述的通讯技术中选出对于系统要求而言性能价格比(cost-effective)最佳的一种。

计算机与外部设备之间的相互连接、与其它计算机之间的连接、以及与通讯系统之间的连接已成为十分重要的设计问题, 但是多年来它一直被人们所忽略。不过许多年来, 设计人员一直在作一些小小的尝试, 将类似于 I/O 那样的数字设备连接起来, 以便为这个富于生命力的领域带来集成化的前景。在这方面有少量的关键性论文和书籍。到目前为止, 仍然有人试图要对通讯网络的整个领域作概括性研究[11]。另有一些研究人员[9, 10, 23 至 30]则试图统一通讯网络的各专业范围中的现行理论, 其中包括信件转接、分布式处理机分类法、总线结构、分布式处理机模型以及线路转接。

数字系统复杂性的发展, 大型通讯网络的出现, I/O 设备的日趋复杂, 以及微处理机技术的发展, 并没有能够对通讯子网理论的发展起到特别的促进作用。到目前为止, 人们对相互连接设计方面的三个特定级别进行了研究, 并取得了进展, 这正是本书所感兴趣的内容。这三个级别是:

1. 系统级[11, 20, 23 至 27, 31 至 39]——主要集中在网络的设计和系统



构成。

2. 通路[9, 11, 25]——包括总线设计, 共享存储器通讯, 以及数据链路。

3. 转接器[3, 6, 8, 19 至 21, 23, 28 至 39]——包括线路转接的设计(分时多路开关和交叉开关)[19]以及信件转接。

本书是为设计师编写的。其目的是为了统一有关通讯子网的设计观点及其对系统的最终影响。主要设计参数将分别在各章中讨论。在每一章中, 我们都列举了一些需要对参数和设计观点进行总体描述的设计方案。一旦有可能, 我们就提出设计理论。在不能提出设计理论的各章, 我们就尽我们所知, 系统地列举各种实用的设计观点、以及可能遇到的主要问题。然而, 本书并不涉及到排队理论, 也并不打算讨论所述设计技术的各个数学模型。

### 参 考 文 献

[ 1 ] E.D.Jensen, "The Influence of Microprocessors on Computer Architecture: Distributed processing," Proceedings of the ACM National Conference 1975:125-128

[ 2 ] K.J.Thurber, "Computer Communication Techniques," COMPSAC IEEE Computer Society, 1978:589-594

[ 3 ] S.R.Kimbleton and G.M.Schneider, "Computer Communications Networks: Approaches, Objectives, and Performance Considerations," Computing Surveys September 1975:129-173.

[ 4 ] K.J.Thurber and H.A.Freeman, "Local Computer Network Architectures," COMPCON, Spring 1979:258-261

[ 5 ] Proceedings of the IEEE, Special issue on packet communication networks, 66(November 1978).

[ 6 ] L.G.Roberts, "The Evolution of Packet Switching," Proceedings of the IEEE 66(November 1978):1307-1313.

[ 7 ] K.J.Thurber, Large, Scale Computer Architecture: Parallel and Associative Processors(Rochelle Park, N.J.:Hayden, 1976).

[ 8 ] W.A.Wulf and C.G.Ball, "C.mmp: A Multi-Mini-Processor," 1972 Fall Joint Computer Conference, 765-777.

[ 9 ] K.J.Thurber et al., "A Systematic Approach to the Design of Digital Busing Structures," Proceedings of the FJCC, 1972:719-740.

[10] K.J.Thurber, "Circuit Switching Technology: A State-of-the-Art Survey," COMPCON, Fall 1978:116-124.

[11] E.D.Jensen, K.J.Thurber, and G.M.Schneider, "A Review of Systematic Methods in Distributed Processor Interconnection," International Communications Conference, 1976:7.17-7.22.

[12] E.D.Jensen and W.E.Bobert, "Partitioning and Assignment of Distributed Processing Software," COMPCON East, 1976.

- [13] H.Lorin, Parallelism in Hardware and Software:Real and Apparent Conccrrency(Englewood Cliffs, N.J.:Prentice-Hall, 1972).
- [14] E.W.Dijkstra,"The Structure of the 'The' Multiprogramming System," Communications of the ACM, May 1968.
- [15] P.B.Hansen, "The Nucleus of a Mult!programming System, "CACM, April 1970.
- [16] L.P.Rivas, H.C.Ludlam, and R.T.Braden,"An Implementation of the MSG Interprocess Communication Protocol, "AD A041630, National Technical Information Center.
- [17] C.G.Bell and A.Newell, Computer Structures:Readings and ExamP-les(New York:McGraw-Hill, 1971).
- [18] E.C.Luczak, "Global Bus Computer Communication Techniques,"Pro-ceedings of the 1978 Computer Networking Symposium, 58-67.
- [19] G.M.Masson, "On Reerrrangeable and Non-blocking Switching Networ-ks.ICC, 1976.
- [20] Digital Equipment Corporation, Distirbuted Systems Handbook(Mayn-ard, Mass.:Digital Press, 1978).
- [21] G.Falk,"A Comparison of Network Architectures: The ARPANET and SNA, "NCC, 1978, PP.755-763.
- [22] G.LeLann, "Distributed Systems-Towards a Fomal Approach,"Proce-edings of the International Federations of Information Processing Societies,1977, PP.155-160.
- [23] F.E.Heart et al."A New Minicomputer/Multiprocessor for the ARPA Network, "NCC 42(1973), PP.529-537.
- [24] G.A.Anderson and E.D.Jensen,"Computer Interconnection:Taxonomy, Characteristics, and Examples, "ACM Computng Surveys, December 1975.
- [25] R.C.Chen, "Bus Communication Systems"(Ph.D.diss.Carnegie-Mell-on, 1974).
- [26] D.P.Sieiwiorek, "Modularity and Multi-Microprocessor Structures," Proceedings of the Seventh Annual Workshop on Microprogramming, October 1974.
- [27] D.P.Siewiorek, "Modularity and Multi-Microprocessor Structures," Infotech Stare-of-the-Art Report on Distributed Processing, 1976.
- [28] M.J.Marcus, "Desingns for Time Slot Iutерchanges,"Proceedings of the Narional Electronics Conference 1970, 812-817.
- [29] N.Pippenger, "On Crossbar Switching Networks,"IEEE Trans on Co-mm, June 1975.
- [30] V.E.Benes, Mathematical Thcory of Connecting Networks and Telep-hone Traffic(New York:Academic Press, 1965).

- [31] R. Rice, "A Project Overview, "COMPCON, 1972.
- [32] R. J. Swan, S. H. Fuller, and D. P. Siewiorek, "Cm\*-A Modular, Multi-microprocessor, "NCC, 1977.
- [33] S. Fuller, D. Siewiorek, and R. Swan, "Computer Modules-An Architecture for a Modular Multi-Microprocessor," Proceedings of the 1975 ACM National Conference,
- [34] Phillip H. Enslow, ed-Multiprocessors and Parallel Processors (New York: John Wiley & Sons, 1974).
- [35] P. H. Enslow, "Multiprocessor Organization-A Survey," ACM Computer Surveys, March 1977, pp. 103-129.
- [36] R. M. Metcalfe and D. R. Boggs. "Ethernet: Distributed Packet Switching for Local Computer Networks, "CACM 19(July 1976), PP. 395-404.
- [37] David J. Farber and Kenneth C. Larson. "The System Architecture of the Distributed Computer System-The Communications System." Proc. Symposium on Computer-Communications Networks and Teletraffic, April 1972 (Brooklyn, N. Y.: Polytechnic Press, 1972).
- [38] Proceedings of the IEEE, Special issue on circuit switching, September 1977.
- [39] Computer, Special issue on circuit switching, June 1979.

## 第二章 各种拓扑概念的通讯结构

### 引 言

到目前为止,笔者曾听说过许多有关分布式处理系统的不同概念,但是它们大多数并没有越出本书等一章所述的简单概念的范围,即分布式系统(原文为 Processor,但第一章中为 system,因此疑为笔误——译注)是这样一种系统,在该系统中,处理机负载系分布在若干处理单元上,这种单元在物理上相互连接,但是允许相互进行数量不等的软件通讯。显然,多处理机和网络都符合这种定义。是否存在对这种系统进行分类的方法,使读者能一眼看出系统的基本结构,从而从感性上认识分布式处理机结构的基本性质呢?答案毫无疑问是肯定的!但是,如果读者本身的文化修养不够,那么即使有这种方法也是无济于事的。

分布式处理机的概念要涉及到许多学科,例如,网络微处理机技术的有效应用,计算机通讯原理等等[1]。有五种分类法和学科论述了分布式处理机的结构;Enslow[2]主张分布式处理机必须具有分布的硬件处理机、分布的数据库以及分布的系统控制;Thurber和Freeman[3]则认为各分布式处理机是紧密相关的,并和计算机局部网络的结构概念交织在一起;对于分布式处理机,Chen[4]持有拓扑学观点;Siewiorek[5]用拓扑结构和通讯方面的概念,对分布式处理机进行了论述,而 Andersen 和 Jensen[6]则从信件通讯的概念出发,考察了分布式处理机的结构。从上述种种观点出发,可以清楚地看出,分布式处理结构可以有许多不同概念,但处理单元之间的通讯从原理上来说是用总线结构(以及 I/O 连接),线路转接、以及报文分组转接来实现的。读者在本章中可以了解系统级结构的概念。通讯结构放在以后几章中讨论。

### 模型基础

Anderson 和 Jensen[6]模型以下列三种简单概念为基础:

信件——在处理机上执行的各进程之间传输的一种信息单位。

通路——信息传输的中间介质,如总线、数据链路等等。

转接器——在发送单元和接收单元之间为信件选路的智能干预装置。

从定义中可以看出,进程 A 和进程 B 之间的信件通讯既可以作为一条通路上的直接通讯,也可以作为通过转接器的间接通讯,在转接器中要发生某些干预,例如信件选路和地址变换等。转接器既可以是集中式的,也可以是非集中式的。这两种决策代表着系统的策略。只要读者理解,这一通讯过程可以在图 1-1 的基础上作进一步分介,以及在低一级中,转接器实际上用于转接某种类型的线路,那么在系统级上就无需对信件转接、报文分组转接或线路转接加以区分。Anderson 和 Jensen 紧接着又建议作出两个战术上的决策——传输结构究竟是专用的还是共享的,以及系统的拓扑结构。图 2-1 和 2-2 中给出了综合分类法,并附有系统结构方面现存的各种重要的系统概念。

## 补充论点

在考察 Anderson 和 Jensen 分类法的过程中, 我们提出一些补充论点供大家参考, 其中包括对原始目标的阐述, 这些论点不是对 Anderson 和 Jensen 分类法的非难, 而是供读者记在心中, 以便全面地理解 Anderson 和 Jensen 分类法的重要性和局限性。

为了深入理解本章的后半部分内容, 对 Anderson 和 Jensen 分类法作一些补充考虑是必要的。大致有下述几个方面:

1. 面向模型的通讯。
2. 模型的属性。
3. 对 I/O、存贮器、软件控制层、处理单元系统接口、以及进程间相互通讯问题缺少考虑。
4. 缺乏对混合系统的讨论。

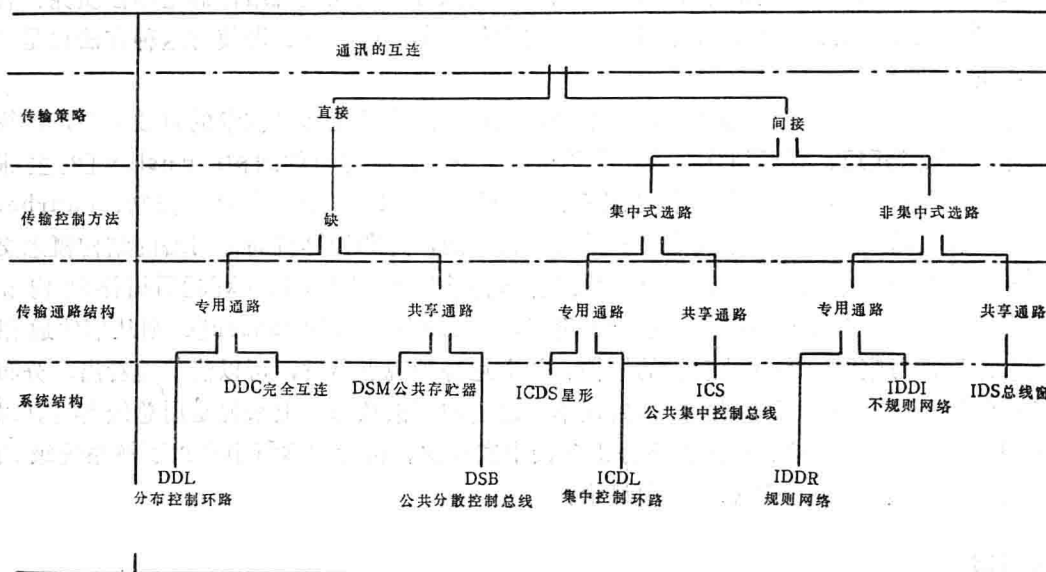


图 2-1 通讯系统模型：Anderson 和 Jensen 分类法

传输策略：直接

通路：专用

拓扑结构：分布控制环路

拓扑结构：完全互连

通路：共享

拓扑结构：公共存贮器

拓扑结构：公共分布控制总线

传输策略：间接

选路：集中选路

通路：专用

拓扑结构：星形

拓扑结构：集中控制环路

通路：共享

拓扑结构：公共集中控制总线

选路：非集中选路

通路：专用

拓扑结构: 规则网络  
拓扑结构: 不规则网络  
通路: 共享  
拓扑结构: 总线窃

图 2-2 Anderson 和 Jensen 分类法轮廓

先考虑第 1 点, 我们假定模型是通过信件进行通讯的。但是它掩盖了这样一个事实, 即信件在总线上的传输不同于信件在转接通讯情况下的传输。读者应该注意, 可以简单地将信件看成是由三个部分组成的一串二进制位, 这三个部分即标题或标识、信体或电文、以及结束标志或终止符。在面向通路和面向转接的系统中, 由于每个信区所允许的规模、所采用的分隔信区的技术, 以及标题和/或结束符为隐式或显式等因素, 使得信件各不相同。例如, 在一个由专用线路和同步处理单元所构成的系统中, 信件可以不需要标题和结束符。另外, 在 Andersen 和 Jensen 模型中没有对线路转接、信件转接或报文分组转接加以区分。只要读者注意到转接器控制着较低级别上的通路或线路, 那么这一假设就还是合理的。这一模型叙述了在最起码级别上所进行的通讯及其实现。在许多情况下, 它掩盖了读者在下几章中将会看到的种种细节。该模型没有讨论迭加(super-Position)的概念, 即两个一般概念可以组合成兼有每个系统优越性的混合系统(第 4 点)。确实, 我们完全可以想象得出这样一个系统, 其中各个主要系统功能(I/O, 存贮器-PE 连接, 后备存贮器层次存贮连接等等)均采用不同的通讯技术。不过 Anderson 和 Jensen 确实曾指出过一些混合系统的例子。

对于第 3 点中指出的许多论点, 如进程间的相互通讯, 我们将予以忽略。在以后几章中, 我们将把 I/O, 控制层次以及 PE-系统接口等问题作为通路设计的一部分。

### 各系统概念的比较

我们选用 Andersen 和 Jensen 所选用的有效性衡量标准来考虑系统级设计问题, 这些标准是:

- (1) 模块性(代价模块性及布局模块性);
- (2) 连接灵活性;
- (3) 容错性;
- (4) 逻辑复杂性;
- (5) 瓶颈问题(即阻塞问题)。

这些概念定义如下:

代价模块性(CM): 增加一台处理单元(PE)所需增加的費用, 在大多数系统中为免费(零代价), 在另一些系统中, 要引入第 $n$ 台 PE, 就要求增加  $n-1$  条通路。

布局模块性(PM): 一台单元的功能或位置所受限制的程度。在大多数系统, 可以方便地加入通路、转接器或 PE, 以便改善性能。在另一些系统中, 例如星形系统, 中央转接器不服从布局模块性。在这种情况下, 增加一个转接器, 就会把结构改变成分类法中的另一种类。

连接灵活性(CF): 它在信件转接系统中尤为重要。它是增加一台 PE 所需化费代价的另一种衡量标准。在增加一台处理机时, 究竟要增加额外的长途连接通路还是增加额外的本机通路, 是衡量连接灵活性的标准之一。这一论点同面向通路有关的模块性方案相似。

容错性(FT): 它是对系统中一次故障所产生的影响的衡量标准。在系统构造时, 它表现为如何按照结构或性能重新构造系统(故障弱化)。具有非集中控制的系统是容错性最好的典型设计。

逻辑(原文为 local, 疑为笔误——译注)复杂性(LC): 它是在一个系统中为信息(发送一封信件)选路所需决策的数量和复杂性的衡量标准。

瓶颈问题(B): 它是由于系统资源要求, 从而使诸如数据链路之类的系统设备饱和或阻塞而产生的系统性能限制。

我们已经讨论了全部十种 Andersen 和 Jensen 的系统概念。

## 设计方案

有几种系统设计方案 Anderson 和 Jensen 并没有作过讨论。这些方案只是隐含在 Anderson 和 Jensen 所讨论的系统中。现在我们列举如下几种论点:

### 1. 系统级

系统中的PE是否必须为同一种类?

系统中转接器是否必须为同一种类(总线、数据链等是否必须为同一种类?)或者转接器可以是不同种类?

通讯线路为双向还是单向?

是否可进行全双工传输?

是否可在一个系统中增加多个通讯设备?

### 2. 处理单元

PE 具有何种类型的缓冲能力?

如果一台 PE 具  $n$  个输入/输出端口, 那么一台 PE 是否可接收  $n$  封信件、发送  $n$  封信件、并同时加以处理?

### 3. 信件

如何检测并处理传送错误?

将信件送往一台不存在的 PE(进程), 将发生什么情况?

信件应答系从 PE<sub>j</sub> 至 PE<sub>i</sub>, 还是只能从源 PE 至目的 PE?

一个系统中是否可同时发送一封信以上信件?

## 系统类型

### 环路结构

环路系统(图 2-3)采用一种直接传输策略, 没有传输控制, 通路为专用[7,8,9]。它们是从数据通讯结构演变而来的代表性系统。在典型的环路系统中, 每个 PE 均与它的两个紧邻连接。从概念上说, 数据通讯联络在这种系统中沿顺时针和/或逆时针方向流动, 但是在大多数已经实现的环路系统中, 总的趋势却是单向性的。系统的工作非常简单——一个处理单元(源)发送一封信件到它的紧邻 PE(目的), 这一步骤一直继续到信件抵达它的终点为止。在抵达终点之前, 信件可能已经经过了许多中间 PE。如果一封件进入某一目的 PE, 而又必须再次发送到另一个 PE, 则该发送 PE 就起了中继站或缓冲器的作用。于是, 各 PE 既要起发送源作用, 又要起目的地作用。在系统结构中作如下选择:

1. 选择选路策略, 根据 PE 的缓冲能力, 在给定时间内仅允许一封信件循环, 或是允许一封以上信件循环。
2. 既允许信件长度固定, 又允许信件长度可变。
3. 允许双向通讯。
4. 允许串行或并行通讯。
5. 允许虚拟信件寻址(发送一封信件到一台逻辑处理机而不是物理处理机, 这就使通讯与处理机数目以及处理机进程—处理机分配无关)。
6. 在一个系统中配备多个环路。
7. 拨出一条环路供某种特定形式的通讯联络专用。
8. 在系统中配备子环路。
9. 配备一组运行速度各不相同的层次环路。

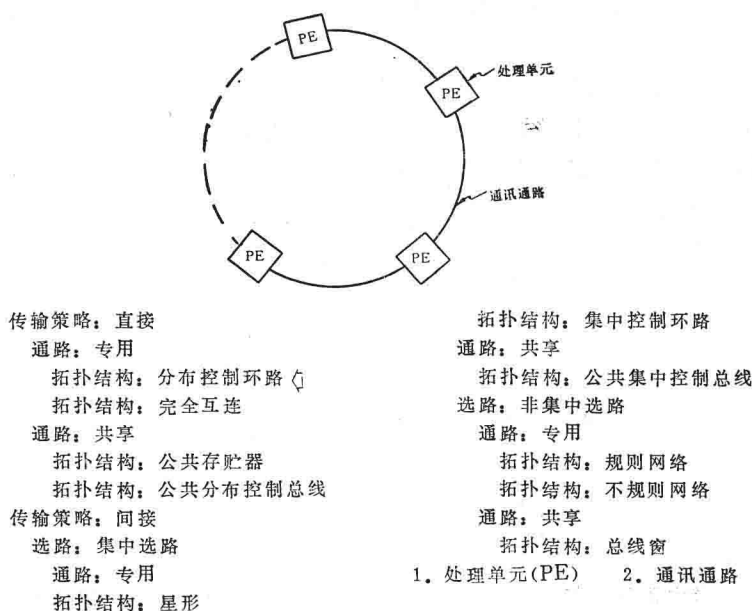


图 2-3 环路系统

环路系统的主要性质如下:

1. CM 很好, 增加一台 PE 很方便, 虽然环路的总延迟要增加; 代价是一条单通路。
2. PM 很好, 对于加入 PE 的种类没有限制。
3. CF 固定。
4. FT 很差, 对单点故障敏感, 这个问题可以通过增加通路和旁路开关来解决。
5. LC 很低, 一台 PE 只需要有识别、接收/缓冲、以及发送信件的能力即可。
6. 潜在的瓶颈可能很坏, 环路的带宽决定了系统是否有瓶颈。

#### 全互连结构

全互连系统(图 2-4)采用直接传输策略, 设有传输控制, 通路为专用[10, 11]。人们有时将现有的计算机相互连接组成系统, 而全互连正是这样演变而来的典型系统。根据定义, 在一个典型的系统中, 所有 PE 均用专用连接连到每一个其它 PE 上。信件在 PE 之间的专用连接线上传输。在作为发送源时, PE 要选择正确的传输通路, 在作为目的地时,



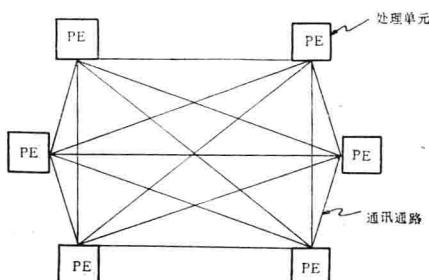
PE 又必须进行仲裁, 以决定在许多可能发来的信件中究竟接收哪一封。作为发送源时, PE 最好能发送多封信件。

系统可作某些选择, 并受某些限制:

- (1) 为了协调性能, 应限制 PE 的发送和接收能力。
- (2) 仅采用本地信件寻址(否则就成了间接传输)。
- (3) 根据链路类型, 可以在两台 PE 间进行双向同时通讯。
- (4) 可具有直接处理 PE-PE 错误的能力。
- (5) 在两台 PE 间具有可变容量数据链路的能力。

全互连结构的主要性质如下:

1. CM 很差, 如果系统原有  $n$  台 PE, 则即使在 PE 端口数允许的情况下, 增加一台 PE 也要增加  $n$  条连线。
2. PM 非常好, 只要 PE 端口数允许, 可以不加限制地加入任何类型的 PE。
3. CF 固定。
4. FT 很好, 当某一链路出现故障时有许多其它道路可供选路, 如果选路是由、或者能够由 PE 承担, 则 PE 出现故障时, 系统还能屏蔽该故障 PE。
5. LC 低, PE 直接将信件发送到另一台 PE。
6. 瓶颈好, 可以为一个特定工作定出通路的带宽和 PE 的数目。



- |                |                        |
|----------------|------------------------|
| 传输策略: 直接       | 拓扑结构: 集中控制环路           |
| 通路: 专用         | 通路: 共享                 |
| 拓扑结构: 分布控制环路   | 拓扑结构: 公共集中控制总线         |
| 拓扑结构: 全互连      | 选路: 非集中选路              |
| 通路: 共享         | 通路: 专用                 |
| 拓扑结构: 公共存储器    | 拓扑结构: 规则网络             |
| 拓扑结构: 公共分布控制总线 | 拓扑结构: 不规则网络            |
| 传输策略: 间接       | 通路: 共享                 |
| 选路: 集中选路       | 拓扑结构: 总线窗              |
| 通路: 专用         | 1. 处理单元(PE)    2. 通讯通路 |
| 拓扑结构: 星形       |                        |

图 2-4 完全系统互连

### 多处理机结构

中央存贮器多处理机(图 2-5)采用直接传输策略, 无传输控制, 传输道路(存贮器)为共享。多处理机是一种很普通的结构, 它是从各 PE 通过共享存贮器进行通讯的系统概念、或要求有大量冗余处理能力的系统演变而来的[12,13,14]。从实际使用的系统来看,