Introduction to Data Communications and Computer Networks

Fred Halsall

Introduction to Data Communications and Computer Networks

Fred Halsall

ADDISON-WESLEY PUBLISHING COMPANY INC.

Wokingham, England · Reading, Massachusetts · Menlo Park, California
Don Mills, Ontario · Amsterdam · Sydney · Singapore · Tokyo · Mexico City · Bogota
Santiago · San Juan

ELECTRONIC SYSTEMS ENGINEERING SERIES

Consulting Editor **E L Dagless**
University of Bristol

Cover design by Sampson/Tyrell Limited.
Photoset direct from the author's text by Quorum Technical Services Ltd.
Printed in Great Britain by R. J. Acford

British Library Cataloguing in Publication Data
Introduction to data communications and computer networks – (Electronic systems engineering.)
1. Computer networks 2. Data transmission systems
I. Halsall, Fred II. Series
001.64'404 TK5105.5

ISBN 0-201-14547-2
ISBN 0-201-14540-5 pbk

Library of Congress Cataloging in Publication Data
Halsall, Fred.
Introduction to data communications and computer networks.
Bibliography: p.
Includes index.
1. Data transmission systems. 2. Computer networks.
3. Local area networks (Computer networks).
4. Computer network protocols. I. Title.
TK5105.H35 1985 001.64'404 85-5974
ISBN 0-201-14547-2
ISBN 0-201-14540-5 (pbk.)

ABCDEF 898765

Preface

The subject of data communications was, until recently, a relatively specialised area of electronic engineering. The dramatic advances in computer technology over the past few years, however, have resulted in an increasing number of information processing systems now being implemented as a linked set of computer-based equipment. The latter include not only computers per se but also a wide range of other types of equipment which incorporate some local intelligence: computer-based visual display terminals, advanced workstations as used in an automated office, intelligent instrumentation equipment as used in industrial process plants, point-of-sale terminals as used in large department stores, and so on. This means that an essential consideration in the design of most forms of computing equipment installed today is the type of data communications facility which is to be used to allow the equipment to communicate with other similar devices. In many instances this necessitates a knowledge not only of the alternative types of data transmission circuits that may be used but also an understanding of the interface requirements to the many different types of computer communication networks which are emerging for this purpose. Data communications and the allied subject of computer networks have thus become essential topics in all modern courses on computer systems design. This book has been written as an introductory text to both these subjects.

Although there are currently a number of text books on the market which cover some of the topics addressed in this book, the majority have been written primarily as reference texts and hence are best suited to practising engineers. Also, much of the material in this field is currently only in the form of standards documents which are intended primarily for system implementors. One of the primary aims of this book, therefore, is to assemble and convert much of this information into a form which is suitable for use by a student or a practising engineer who wishes to gain a general understanding of the different types of data communication networks and the associated devices and techniques which are used for connecting equipment to these networks.

The book is based on a course of lectures given by the author to both electronics engineering and computer science students and hence care has been taken to avoid any significant prerequisite knowledge in either subject. The amount of material covered in the book is sufficient for a course of lectures which runs over a full academic year but it has been structured so that

it is possible for a lecturer to select just a number of chapters and still present a coherent set of lectures. Although the book has been written primarily as an undergraduate text, the avoidance of any significant prerequisite material means that it is also suitable for students on diploma and higher technician courses.

Because of the diverse range of applications of information processing systems, it is perhaps not surprising that there is also a wide range of data communications equipment and networks which are in use to link the various equipments together. For clarity, therefore, the different types of data communications facility and the associated hardware and software which are required for use with them are treated separately in different chapters. Whenever possible, however, the common underlying principles of the different types of system are emphasised.

In Chapter 1, a brief account of the historical evolution of information processing systems is presented. This is then followed in Chapters 2 and 3 with a description of the basic devices and techniques which are used for the reliable (error free and without replication) transmission of data between two pieces of equipment connected by a point-to-point data link. These three chapters, therefore, are considered to be essential reading and hence should be included in any course of lectures on this subject. Each of the following chapters is then reasonably self contained and it is possible to include or omit topics to suit the type of course being offered.

Chapter 4 is concerned with the properties of the electrical interface between the interconnected items of equipment and the different types of physical transmission media. The various internationally agreed standards which have been defined for this purpose are presented and, although care has been taken to avoid any specific previous knowledge of electronics, this topic has been deliberately covered in a separate chapter so that, if required, it can readily be omitted without any loss of continuity.

Chapter 5 is concerned with the different types of communications equipment which is in use in terminal-oriented distributed computing systems. Descriptions are presented of both the function and mode of operation of these types of equipment and also the alternative communications protocols which are in use in such systems.

Chapter 6 is concerned with descriptions of the various communication protocols which have been defined to enable computers from different manufacturers to freely exchange information. It is specifically concerned with the four highest protocol layers in the International Standards Organisation (ISO) Reference Model for open systems interconnection. It thus includes descriptions of the function and operation of the Application Layer, the Presentation Layer, the Session Layer, and the Transport Layer protocols. The descriptions presented do not strictly adhere to the standards – indeed some of the protocols are not yet finalised – but rather, are intended to give the reader a general understanding of the role of the various protocol layers in the Reference Model.

The lowest three protocol layers in the ISO Reference Model for open

systems interconnection vary according to the type of data communications network which is being used to connect the various computers together. Thus Chapter 7 describes the function of the three lowest protocol layers which are used to interface a computer to a public data network and Chapter 8 with those required to interface a computer to a local area data network. Although there are currently a large number of different types of local area network in existence, only those which are currently in the various standards documents are described. Also, in addition to the protocols used in these networks, both chapters describe the function and operation of the various items of additional hardware and software which are required to interface a computer to each type of network.

I would like to take this opportunity to express my sincere thanks, firstly to Dr. Keith Bennett at the University of Keele for reviewing the draft manuscript and for his many helpful suggestions; secondly, to Christine Thornton-Clough for typing the manuscript and making numerous corrections and alterations without any word of criticism; and finally to my wife Rhiannon and children Lisa and Richard for their patience and understanding whilst I was writing the book.

Fred Halsall,
University of Sussex

Contents

Chapter 1 Distributed System Architectures

Objectives:
When you have completed studying the material in this chapter you should be able to:

- describe the historical development of the different types of distributed computing system;
- draw schematic diagrams to illustrate the architecture of the different types of distributed system;
- appreciate that irrespective of the application of the system, there is only a limited number of different types of data network that may be used to provide the underlying data communication services;
- understand that there is a degree of commonality to the structure and function of the communications software used within the various interconnected equipments which make up the system, which is independent of both the application of the system and the type of data communications network being used.

1.1 INTRODUCTION

Distributed computing systems are concerned with the processing and the communication of information between distributed communities of electronic digital equipment. In general, the various types of equipment are referred to as *data terminal equipments* or simply *DTEs*. These include not only computers but also a wide range of other devices: visual display terminals, for example, computer-based office workstations, intelligent instrumentation equipment for industrial process control, point-of-sale terminals as used in large department stores, microprocessor-based domestic electricity meters designed for remote reading, and many others.

This wide range of devices means that there are many different types of distributed systems. For example, a system may be made up of a large community of visual display terminals physically distributed over a wide geographical area and communicating with a large centralised computing complex. Alternatively, a system may comprise a number of computer-based office workstations physically distributed around a single block of offices

providing, for example, word processing functions and access to various shared resources – printers, copiers, file systems etc.

Although the overall processing functions performed in the different types of distributed system may vary considerably from one application to another, there is only a limited number of types of data network that may be used to provide the underlying data communication services for the exchange of information between the various equipments. Moreover, the advent of international standards for both the structure and function of much of the communications software, which is needed within each piece of equipment to achieve the reliable exchange of information over these networks, means that there is a degree of commonality to the communications software within each equipment which is independent of the particular application of the overall system. When designing the data communications facilities to be used with any form of distributed system, therefore, it is necessary to have a working knowledge firstly of the different types of data communication networks which are available and their corresponding modes of operation and application areas, and, secondly, an understanding of the various international standards which have now been established to aid the use of these networks. This book is specifically concerned with these aspects of distributed computing systems.

1.2 HISTORICAL EVOLUTION

The evolution of distributed computing systems can perhaps best be traced by following the development of the computing resources used within any large organisation. The earliest commercially available computers that were used were characterised by expensive hardware and relatively primitive software. Typically, an organisation would purchase a single computer system which would then be centrally located in a large, air-conditioned room. It would consist of a central processing unit (CPU) with a limited quantity of primary (RAM) memory, some secondary (tape or drum) storage, a printer, a punched-card reader and an operator console. Users normally prepared their programs and data *off-line* on a card punch located in a different room and the operator would then load and run the prepared programs sequentially. This type of early computer system is shown in diagrammatic form in Fig. 1.1.

As computer technology and its operating software advanced, fast secondary storage – large magnetic drums and later discs – and multiprogramming operating systems were developed making it possible to *time-share* the central processing unit between a number of active programs (processes) thereby allowing multiple users to use the computer simultaneously. Each user was allocated a separate terminal which then allowed multiple users to access simultaneously stored information and run programs *interactively* at the terminals. The latter were normally electromechanical teletypewriters (TTYs) similar to those which were already in use in the international telex networks. They were designed, therefore, to transmit and receive data over long distances and operated in a *serial mode*.

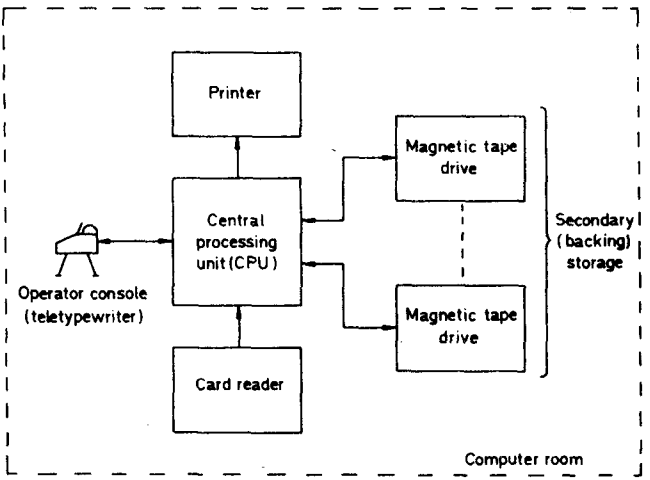

Fig. 1.1 Typical early computer system

To exploit these developments, the computers used within organisations were upgraded to support, say, five or more such terminals. The computers then became known as *multi-access* systems providing *on-line* access to stored data. Initially the terminals were all located close to the main computer complex but, because of their basic mode of operation, it soon became common practice to distribute the terminals around the organisation, firstly locally in different offices and later, with the aid of the ubiquitous switched telephone network and *modems*, nationally over wide geographical areas. A typical computer system in operation at that time, therefore, is as shown in Fig. 1.2.

The use of the switched telephone network as the basic data communications medium meant that communication line costs were not insignificant and indeed soon became a substantial proportion of the system operating costs. In order to minimise these costs, therefore, devices such as *terminal multiplexers* and *cluster controllers* were introduced. Essentially, these allow a single communication line – often permanently leased from the public telecommunications authorities – to be shared between a number of simultaneous users all located, for example, at the same remote site. In addition, the increasing level of usage of the computer within the organisation soon gave rise to systems containing many hundreds of terminals with the effect that the central computer could no longer cope with the processing overheads associated with servicing the various communication lines on top of its normal processing functions. This in turn gave rise to the *front-end processor (FEP)* which essentially off-loaded the processing overheads associated with the various communication lines from the central machine. This effectively signalled the beginning of distributed computing sytems and a

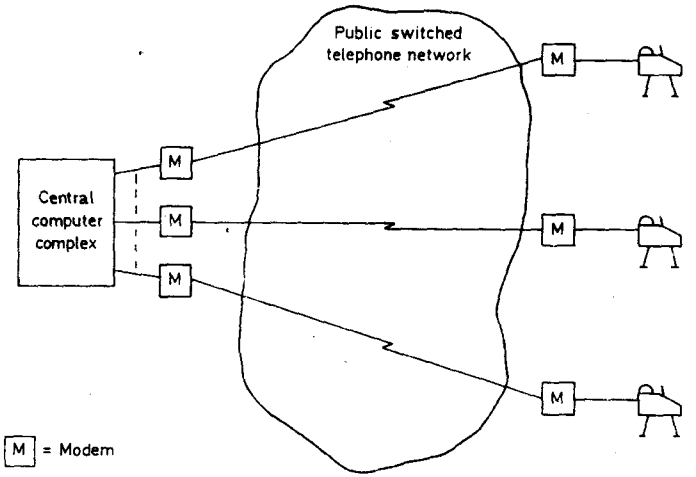

Fig. 1.2 Early terminal-oriented distributed system

large but typical system in use at this time is as shown in Fig. 1.3. This type of distributed system is still in widespread use today and aspects of this type of system are described in Chapter 5.

1.3 PRIVATE COMPUTER COMMUNICATION NETWORKS

The structure shown in Fig. 1.3 was particularly prevalent in large organisations such as the major clearing banks and airlines which normally held large quantities of information at a single central site. The distributed community of users then accessed and updated this information using the communication facilities outlined. In many organisations, however, it was not necessary to hold all information centrally and hence it soon became common place for an organisation to have a number of autonomous computer systems located at different sites around the country. Typically, these provided a local computing function but there was often a requirement for them to communicate with each other both to share resources – hardware and software – and also to exchange information. The requirements for the communications facility to meet this type of interconnection, however, are in many ways more demanding than a terminal-based system since instead of all communications being directed to a single known destination as before, a more flexible, effectively switched, communications facility is required.

The limited data capacity of a communications link established using the switched telephone network and modems meant that a conceptually different approach to providing the necessary data communications services was

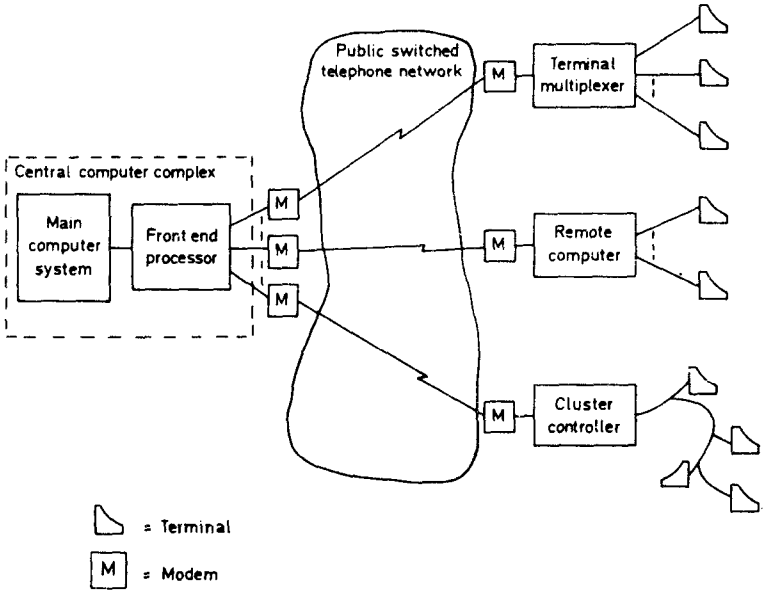

Fig. 1.3 Large terminal-oriented distributed system

required and it was at this point that it became economically more attractive to provide a separate autonomous *data communications subnetwork* to provide the required data communication services. The requirements for such networks were in many ways similar to those provided by the normal telex network. The latter operates in a *message store-and-forward* mode which, as will be expanded in later chapters, is also ideally suited to the sporadic nature of communications between computers and other computer-related equipment.

With the telex network the internal message units used may be long with the effect that the response time of such networks, that is the time delay between a message entering the network and subsequently leaving the network, can be degraded whilst a number of long messages are in transit. In order to overcome this limitation, therefore, communication subnetworks intended for carrying computer data normally operate using a smaller maximum message unit known as a *packet*. The resulting communications subnetwork is then said to operate using *packet-switching* or in a *packet store-and-forward mode*. This type of distributed system is shown diagrammatically in Fig. 1.4 and, because the interconnected computers are normally physically distributed over a wide geographical area, it is also known as a *wide area computer communication network* or *WAN*.

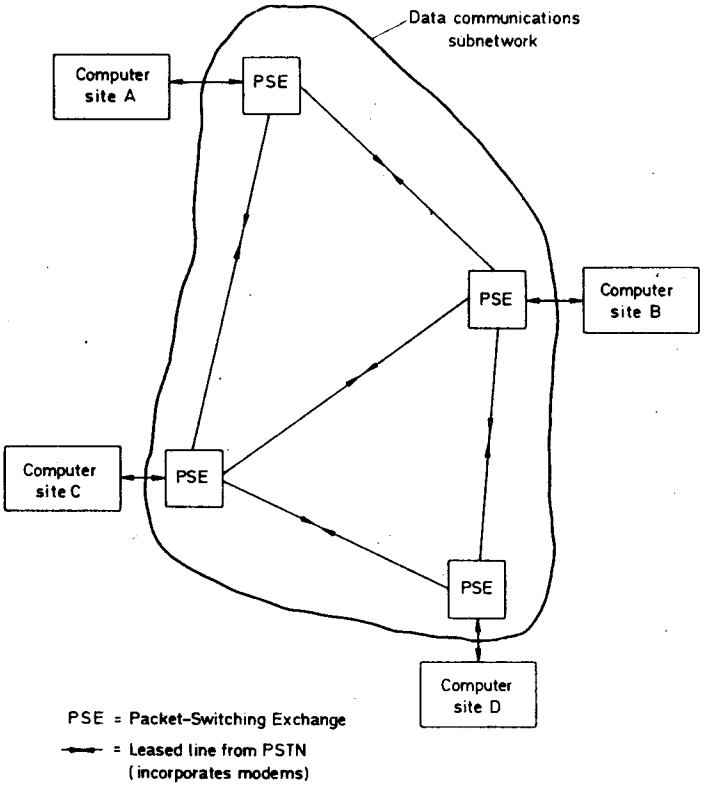

Fig. 1.4 Early private computer communications network

1.4 PUBLIC SWITCHED DATA NETWORKS

Initially, organisations implemented their own private nationwide data communication subnetworks using communication lines leased from the public telephone authorities and their own proprietory switching equipment. The larger computer manufacturers then produced suitable communications software packages to enable their computers to communicate and exchange data using these networks. With time, however, as the impact of computer technology on the operation of organisations grew, the need arose for a computer in one organisation – and hence from one manufacturer – to communicate with a possibly different computer in another organisation; for example, to transfer funds from one bank computer to another.

It was at this point that the *public telecommunications authorities* in a number of countries – generally referred to as *PTTs* for Post, Telephone and Telecommunications – accepted that a *public switched data network (PSDN)* analogous to the normal *public switched telephone network (PSTN)* was justified. Moreover, since it was intended that this should be used to provide a

communications facility for the interconnection of a possibly large number of different computers from a range of manufacturers, the definition of agreed interface standards became all important.

After much discussion in various standards committees, firstly at national and later at international level, a set of internationally agreed standard protocols was defined, firstly for interfacing and controlling the flow of information between the various data terminal equipments – computers and terminals – and the PSDN, and later to control the exchange of information between the two communicating items of equipment. An increasing number of public networks have now been implemented which support these standards. This type of distributed system is shown diagrammatically in Fig. 1.5 and some of the different types of PSDN are described in Chapter 7 of the book. Also, the so-called higher level protocol standards which have been

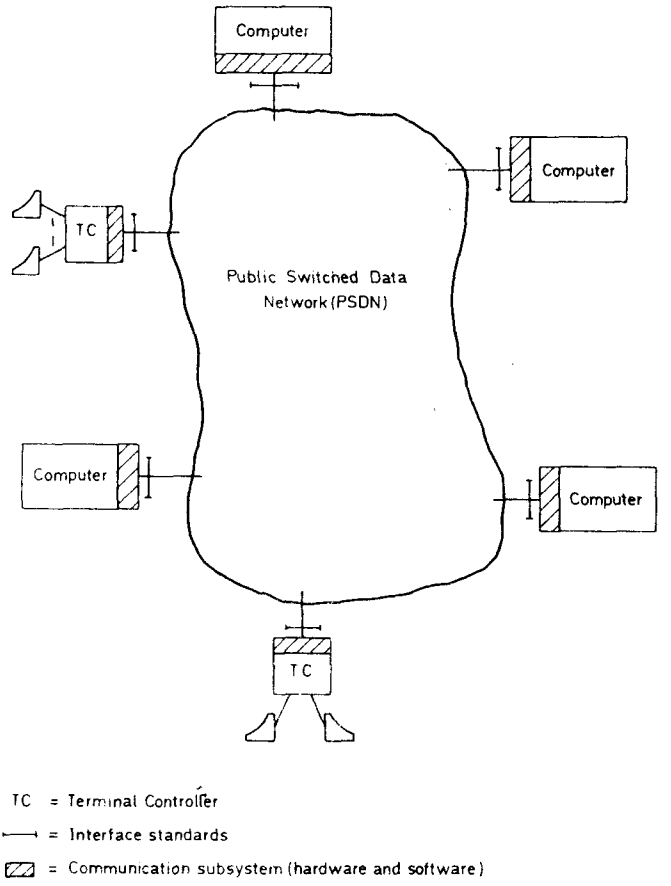

Fig. 1.5 PSDN-based distributed system

defined to control the exchange of information between two pieces of equipment using such networks are discussed in Chapter 6.

1.5 LOCAL AREA NETWORKS

Although the computing resources in an organisation steadily increased with developing technology, the advent of the microprocessor and the associated advances in integrated circuit technology meant that in addition to the systems just outlined, it soon became common to find a multiplicity of different computer-based devices physically located within the same building or block of offices; for example, a community of intelligent computer-based workstations performing word processing and other related functions. Although the availability of processing power within each of these systems means that many computing tasks can be carried out locally, in many situations there is often a requirement for these systems to communicate with each other; for example, to exchange (electronic) mail or to access an expensive shared resource such as a letter-quality printer or file system.

Since the linked computing devices in this type of network are physically located close to one another, the communication facility provided is referred to as a *local area data network* or simply a *LAN*. Although the different types of distributed system outlined in the earlier sections are now relatively widespread and established, this type of distributed system is relatively new. It is perhaps not surprising, therefore, that many of the current (1985) designs of LAN show significant differences from one another, both in terms of their topology and their mode of operation. In order to avoid a plethora of different network protocols being adopted, therefore, the standards bodies have responded swiftly and indeed a range of internationally agreed standards have now been defined for use with this type of network. Moreover, the structure of these standards is in keeping with the structure of the standards which have been defined for use with wide area networks, and it is intended that the same higher-level protocols will be used with both types of network. An example of a small but typical LAN-based distributed system is shown in Fig. 1.6 and the different types of LAN and their associated communication protocols are discussed in Chapter 8.

1.6 SUMMARY

It can be concluded from the above paragraphs that there is a wide range of different types of distributed system in operation today, each of which has been designed to meet a different set of application requirements. Indeed, when designing any form of distributed system, it is important to analyse the particular set of application requirements before deciding on the most suitable form of data communication system to be used. It is for this reason that the book has been organised into a number of distinct sections each devoted to a particular type of distributed system. It is felt that in this way the reader can more readily appreciate the reasons for the use of the specific types of network and protocols adopted.

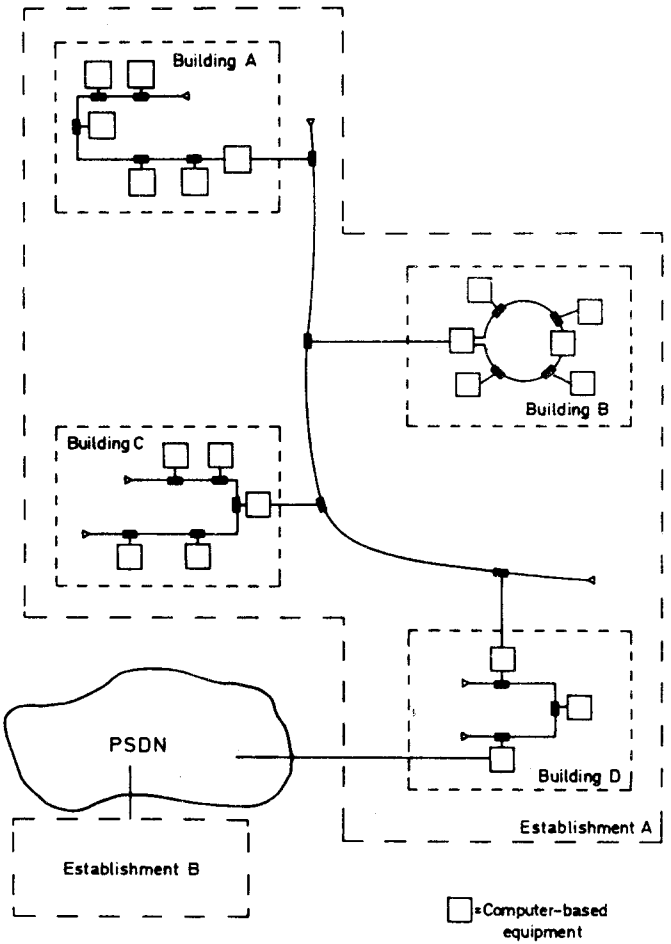

Fig. 1.6 LAN-based distributed system

As has been indicated, the need for data communications has evolved almost from the earliest days of computing. However, although the reader of this book is likely to be aware of the basic terminology and devices associated with computers themselves – bits and bytes, gates and highways, BASIC and Pascal, etc. – equally, there is a fundamental set of techniques and terminology associated with data communications and these are often less well understood. Before commencing with descriptions of the different types of distributed system and network types, therefore, the first two chapters of the book are devoted to a review of the fundamentals and terminology of data communications on which all forms of distributed system are based.