Performance Analysis of Local Computer Networks

Joseph L. Hammond

Peter J.P. O'Reilly

Performance Analysis of Local Computer Networks

Joseph L. Hammond
Clemson University

Peter J.P. O'Reilly
GTE Laboratories, Inc.

Addison-Wesley Publishing Company

READING, MASSACHUSETTS MENLO PARK, CALIFORNIA DON MILLS, ONTARIO
WOKINGHAM, ENGLAND AMSTERDAM SYDNEY SINGAPORE TOKYO MADRID
BOGOTÁ SANTIAGO SAN JUAN

Library of Congress Cataloging-in-Publication Data

Hammond, Joseph L.
Performance analysis of local computer networks.

1. Local area networks (Computer networks)
I. O'Reilly, Peter J. P. II. Title.
TK5105.7.H36 1986 004.6'8 85-13526
ISBN 0-201-11530-1

ABCDEFGHIJ-MA-89876

Preface

This text covers the field of local computer networks with emphasis on performance analysis. In this preface we make a few comments on the evolution and final selection of topics, their relation to the general area of computer networks, and the intended uses of the book.

Computer Networks

Computer networks, or computer–communication networks, is an area that focuses on the study of communication structures—interfaces, connecting media, access and routing algorithms, flow and error control, and related topics. These structures make it possible for physically separated data processing devices to exchange data. The data processing devices include such items as mainframe computers, personal computers, terminals, mass storage units and displays. As the field has developed, it has become apparent that design issues for computer networks are determined to a large extent by the physical size of the area covered by the network. Thus "wide area" networks, which in the extreme span the globe, require different communication structures from "local" networks, which are restricted in extent to one building or the area of a college campus. This text is concerned with the study of local networks.

As with any significant topic, there are many aspects to the design of local computer networks. The more important of these are: communication hardware, protocols, hardware and software for interfacing equipment, access algorithms, and network performance. The major emphasis in this text is on the analysis of network performance. The development, however, is such that significant attention is also given to the other important issues in the design process.

Evolution of Text Material

The material for the text was put together while the authors were with the Georgia Institute of Technology. The evolution of the material to a major extent coincided with the maturing of local area networks as a distinct area.

The idea, first for a course and then for a textbook, originated in the late seventies while one of the authors (JLH) routinely taught a course covering wide area networks. It seemed possible to us to treat in a unified fashion most of the mathematical performance models that were being proposed for local area networks. We also felt that the topic of local networks would not only provide a study area of manageable size, but one of increasing general interest.

The level and emphasis of first the course and then the text were based on the following considerations. Local networks are extended forms of communication systems and thus logically fit with communication courses in an electrical engineering curriculum. Electrical engineering students, especially those in communications, are accustomed to a quantitative approach and to computing system performance. Thus an emphasis on performance analysis of local networks seemed logical, with details of hardware and software provided as essential introductory material. It can be noted parenthetically that, even though the text has evolved in an electrical engineering context, it is not tied to an electrical engineering background or prerequisites. Thus the book should be useful for computer scientists and will complement alternative approaches such as the layered approach emphasizing network architecture and protocols.

Since mathematical models for performance of local networks are at best approximate, it did not seem reasonable to us to develop the models with excessive mathematical rigor. Thus we decided that an undergraduate course in probability was all that was needed as a mathematical prerequisite.

The evolving material was taught in a senior electrical engineering course over several years, first as a special topic course and then as an approved senior elective. The book was developed as an extension of the notes used for these courses.

Organization

The book has eleven chapters, the contents of which are reviewed in Chapter 1. The material can be organized for different courses of varying length: parts of Chapter 2 as a review, Chapters 3 and 5–9 comprise a basic package on performance analysis with essential background. This selection of material can be covered in one quarter and is the basis of the Georgia Tech senior elective.

For a one semester course, the material on PBX local networks in Chapter 4 and/or the material on protocols in Chapter 10 can be added. The bus networks discussed in Chapter 11 can be included with either of the above selections as time permits.

As still another alternative, the material in Chapter 2 can be covered in depth, with some elaboration, to provide an introductory course on data communication. If this route is taken, most of the remaining chapters of the book can be covered in a single quarter or semester.

Objectives of the Text and Required Prerequisites

The book is intended for practicing engineers and for students in electrical engineering and computer science at colleges and universities. The level of the presentation is appropriate for seniors or first year graduate students. Mastery of this material should prepare the reader to carry out design studies through performance analysis and to understand the technical literature in the area.

The text is meant to be quantitative but not rigorous. Every effort has been made to use as little mathematical detail as possible consistent with the study of performance analysis. Emphasis is placed on interpreting and using performance equations rather than on proofs.

The mathematical background required is an undergraduate course in probability. It is preferred that students take a prior course in data/digital communications; however, basic material from this area is given in the first part of the book. A course on general computer networks is not required but clearly would add to the student's perspective in the area. No specific computer languages, information on digital circuit design, or knowledge of operating systems is required.

Most of the chapters of the book contain problems for solution, which may be used as part of a formal course or for self-study.

Acknowledgments

As noted above, the contents of this book have evolved over a number of years through contacts with students and with the developing technical literature. The influence of the latter in defining and structuring the technical content and of the former in shaping the material, into what we hope is an understandable form are gratefully acknowledged.

The text material was reviewed by several persons whose comments contributed to its development. We are especially indebted to John Spragins for a careful review of the complete manuscript and many helpful suggestions. Thanks are also due to Suzie Roe for her encouragement and excellent typing.

Acknowledgment is made of support from a National Science Foundation Local Course Improvement grant (SER–8160975) under which some of the course development took place.

JLH
POR

Contents

PREFACE

1 INTRODUCTION 1

1.1 A Perspective on Computer Communication Networks 1
1.2 Local Computer Networks 3
1.3 Local Network Design and Performance 5
1.4 Outline of Book 6
1.5 Summary 6
References 7

2 DATA COMMUNICATION 9

2.1 Introduction 9
2.2 Digital Communication 10
2.2.1 Signals for Digital Transmission
2.2.2 Bandwidth of Signals and Systems
2.2.3 Modulation
2.2.4 Transmission of Digital Information
2.2.5 Data Rates—Bandwidth/Data Rate Tradeoffs
2.2.6 Signaling Schemes for Local Area Networks
2.2.7 Data Transmission
Asynchronous Data Transmission
Synchronous Data Transmission
2.2.8 Communication Control Procedures
2.3 Error Control 42
2.3.1 Introduction
2.3.2 Parity Checks
2.3.3 Cyclic Redundancy Checks
Systematic Representations of Codes
Implementation of Encoders and Decoders
Properties of Error Detecting Codes
2.3.4 ARQ Schemes

2.4 Switching and Multiplexing 66
2.4.1 Introduction
2.4.2 Switching
2.4.3 Multiplexing
Frequency-division Multiplex
Synchronous Time-division Multiplex
Asynchronous Time-division Multiplexing
2.4.4 Broadband Systems
2.5 Summary 76
References 77
Problems and Exercises 78

3 DATA FLOW IN NETWORKS AND QUEUES 81

3.1 Introduction 81
3.2 Steady-state Data Flow 82
3.3 A Class of Queueing Models 90
3.3.1 M/G/1 Model
3.3.2 M/M/1 Model
3.3.3 M/D/1 Model
3.3.4 General Comments
3.4 Priority Queues 98
3.5 Networks of Queues 105
3.6 Summary 106
References 107
Problems and Exercises 107

4 CIRCUIT-SWITCHED LOCAL NETWORKS 111

4.1 Introduction 111
4.2 Digital PBXs 112
4.3 Digital Switches 115
4.3.1 Space-division Switching
4.3.2 Time-division Switching
4.3.3 Multiple-stage Switches
4.4 Blocking Analysis 123
4.4.1 Transition-rate Diagrams
4.4.2 Loss Systems
4.4.3 Delay Systems
4.5 Examples of PBX System Sizing 135
4.5.1 Sizing of PBXs
4.5.2 Blocking Calculations
4.6 Summary 138
References 139
Problems and Exercises 139

5 INTRODUCTION TO LOCAL COMPUTER NETWORKS 141

5.1 Definition and Functions 141
5.2 Network Structure and Topology 145
5.3 Channel Access Techniques 148
5.4 Network Traffic 150
5.5 Performance Measures 154
5.5.1 User-level Criteria
5.5.2 Network-level Criteria
5.5.3 Specific Performance Criteria Definition and Notation
Throughput
Average Transfer Delay
5.6 Examples of Commercial and Experimental Local Networks 159
5.6.1 Fixed Assignment: Amdax CableNet
5.6.2 Random Access: Xerox Fibernet and MITRE Cablenet
5.6.3 Demand Assignment with Central Control: Airline Reservations Systems (PARS), Stock Quotation System (NASDAQ)
5.6.4 Demand Assignment with Distributed Control: Distributed Loop Computer Networks (DLCN)
5.6.5 Adaptive Assignment: The Adaptive Tree Walk Protocol
5.7 Summary 166
References 166
Problems and Exercises 167

6 PERFORMANCE OF BASIC ACCESS PROTOCOLS 169

6.1 Introduction 169
6.2 Fixed Assignment Access Methods 170
6.3 The ALOHA Random Access Procedure 175
6.4 Idealized Central Control 181
6.5 Comparison of Basic Capacity Assignment Strategies 185
6.6 Summary 189
References 189
Problems and Exercises 189

7 POLLING NETWORKS 193

7.1 Introduction 193
7.2 Operation of Polling Strategies 196
7.3 Performance Analysis 202
7.3.1 Average Cycle Time
7.3.2 Delay Analysis
7.3.3 Average Number Stored in a Station Buffer
7.4 Performance of Polling Networks 211
7.5 Adaptive Polling 216

7.6 Summary 218
References 219
Problems and Exercises 219

8 RING NETWORKS 223

8.1 Introduction 223
8.2 Token Rings 225
8.2.1 Basic Operation
8.2.2 Delay Analysis
Multiple-token Operation
Single-token Operation
Single-packet Operation
8.2.3 Performance Studies
8.3 Slotted Rings 244
8.3.1 Basic Operation
8.3.2 Delay Analysis
8.3.3 Performance Studies
8.4 Register Insertion Rings 255
8.4.1 Basic Operation
8.4.2 Delay Analysis
8.4.3 Performance Studies
8.5 Comments on Comparative Performance of Ring Networks 270
8.6 Summary 273
References 273
Problems and Exercises 274

9 RANDOM ACCESS NETWORKS 279

9.1 Introduction 279
9.2 Basic Properties of Random Access Networks 282
9.3 Slotted ALOHA 286
9.3.1 Throughput Analysis
9.3.2 Delay Analysis
9.3.3 Stability Considerations
9.4 Protocols That Use Carrier Sensing 304
9.4.1 Persistent and Nonpersistent CSMA Protocols
9.4.2 Throughput Analysis of Nonpersistent CSMA
9.4.3 Throughput Analysis for Slotted Nonpersistent CSMA
9.4.4 Additional CSMA Results and Performance Studies
9.5 Protocols That Use Carrier Sensing with Collision Detection 320
9.5.1 Basic Operation
9.5.2 Throughput Analysis for CSMA/CD
9.5.3 Performance Studies for CSMA/CD

9.6 Summary 339
Appendix: Determination of the Average Number of Retransmissions for Slotted ALOHA 340
References 343
Problems and Exercises 344

10 PROTOCOLS AND NETWORK ARCHITECTURE 347

10.1 Introduction 347
10.2 The ISO Reference Model 350
10.2.1 The Seven ISO Layers
10.2.2 Types of Service and Congestion Control
10.3 IEEE 802 Standards for Local Area Networks 355
10.4 Data-link Protocols 358
10.4.1 Basic Data-link Protocols
10.4.2 Logical Link Control Protocols
10.5 Local Network Architectures 366
10.5.1 Intel Local Network Architecture
10.5.2 Local Network Architecture for a Token Ring
10.5.3 Sytek Inc. LocalNet Architecture
10.6 Internetworking 375
10.7 Summary 378
References 379
Problems and Exercises 379

11 IEEE 802 BUS NETWORKS 381

11.1 Introduction 381
11.2 Ethernet 381
11.2.1 The Ethernet Specification
11.2.2 Performance Studies for Ethernet
Results of Measurement and Simulation
Analytic Results
Concluding Comments on Performance of Ethernet
11.3 Token Bus 390
11.3.1 IEEE 802 Draft Standards for Token Busses
11.3.2 Performance Studies for the Token Bus
11.4 Comparison of Bus Access Methods 397
11.5 Summary 401
References 401

INDEX 403

CHAPTER 1

1.1 A Perspective on Computer Communication Networks
1.2 Local Computer Networks
1.3 Local Network Design and Performance
1.4 Outline of Book Contents
1.5 Summary
References

Introduction

1.1 A Perspective on Computer Communication Networks

In the early days of computing, intelligent resources were localized in a single main frame. A few peripherals, such as printers, card readers, or terminals, might be connected to the main frame facilities, but these devices were close by and operated as secondary slave facilities.

As the computing discipline matured and equipment evolved, facilities spread out. More terminals and other peripherals were connected at increasingly greater distances from the main frame. Intelligent devices were moved from the main frame to remote locations, and the concept of computer communication networks came into being.

The term *computer network* is somewhat ambiguous in current usage when it is used synonymously with *distributed system*. There seems to be agreement that a computer network is an interconnection of autonomous computers that are capable of exchanging information. When the independent computers act together to form a distributed system, however, several competing definitions come into play. The focus of this text is not on distributed computing. Thus Tanenbaum's view that "a distributed system is a special case of a network" [1] seems reasonable, and allows the focus to be placed on what occurs in the communication medium that couples the computers and its interfaces, i.e., the network itself. This text focuses on the latter aspect of computer communication networks in the special context of local networks. *Computer communication network* is used in the section heading to emphasize that this text focuses on issues that concern the network as it provides for communication between its attached devices. This term is awkward, however, and the more usual term *computer network* is used throughout this text.

For purposes of this text, a computer network can be viewed as a facility that makes possible communication between computers and other devices. Major components of the network include the connecting links, the interface between devices and the network, and protocols, which are rules for managing the network resources.

Computer networks can be classified in a variety of ways that include their geographical extent. On this basis there are wide area, or long-haul, networks and local, or local area, networks. Long-haul networks extend over wide geographical areas, which in some cases are literally global in extent. Local networks, on the other hand, cover only short distances, usually not exceeding a few miles.

The differences in geographical extent account for significant differences in the design of long-haul and local networks because of the different requirements for the node-to-node links. For local networks it is possible for a single organization to purchase and install high-speed, low-noise links. On the other hand, long-haul networks require the use of shared links for which large bit rates are expensive. Furthermore, links for long networks can be noisy.

For a number of reasons, which include the need to make optimum use of expensive communication links, long-haul networks are structured with irregular placement of the nodes. Modern long-haul networks make use of store-and-forward packet switching, which means that packets of data proceed through the network much like automobiles move over a freeway system. At each node, addresses carried with each packet must be read, and the packets must be routed toward their ultimate destination.

The manner of operation just described causes long-haul networks to operate at relatively low bit rates. Thousands to tens of thousands of bits per second are typical, although satellite links can operate in the megabit range. The long distances usual for the links make propagation time over the links a significant factor.

Representative design issues for long-haul networks include:

- Bit rate capacity assignments for the connecting links;
- Choice of network geometry;
- Design of routing algorithms;
- Congestion control; and
- Design of network architectures, including protocol hierarchies.

Because local computer networks can be designed around a high-speed, low-noise connecting link, these networks operate quite differently from long-haul networks. Local computer networks are usually designed so that all nodes are connected by a single high-speed shared channel that typically supports data rates in the range of 500 kilobits per second (kbps) to 50 megabits per second (Mbps).

Data is packetized, as in long-haul networks, but for local networks the packets all traverse the same channel, and normally the path of a given packet carries it past all nodes on the network. Thus, although addressing information is still required, routing is unnecessary. Several geometries can be used for the shared channel, but choice of the geometry is not a design issue per se. Congestion control and design of network architectures and protocols are issues for local networks as well as for their long-haul cousins.

The design of either type of computer network could be approached on a top-down basis by starting with the network architecture and the protocol hierarchy. Tanenbaum [1], for example, structures his book on computer networks to parallel the seven layers that have been identified for the protocol hierarchy (see Chapter 10 for a discussion of these layers).

1.2 Local Computer Networks

The previous subsection notes that local networks extend over a limited geographical area and are structured around a high-speed channel. Networks of this type have been developed in response to several clear-cut needs.

As hardware costs for intelligent devices have decreased, these devices have been installed in more locations within a single facility, such as an office building or a university campus. It is still expensive, however, to duplicate line printers, disk storage, and other such devices, and, even if it were not, such devices are not fully utilized when they are installed with the intelligent devices at each workstation. Local networks are used as one way to share these expensive resources.

A second motivation for local networks is the need to share data resources. Central data files are costly to duplicate and expensive to maintain. It is often expedient to maintain central files or data banks that are shared by devices at workstations throughout a local computer network.

In addition to the two primary justifications for local networks, there are a number of secondary influences. The distributed resources of a local network provide redundancy for many devices and thus backup in the event of failures. Many local networks provide speed and code conversion, which enables equipment from different manufacturers to be connected easily without expensive special purpose interfaces. Another point is that some networks also provide for local editing of code.

Computer networks were introduced as the coupling of computers. From the preceding discussion, it should be clear that local computer networks are used to couple many types of devices, not limited to computers. Examples of equipment that can be coupled to local computer networks are:

- Computers;
- Word processors;
- Terminals;
- Graphics terminals;

- Local storage;
- Line printers;
- Telephones; and
- Data files.

As in any new area, terminology and definitions are somewhat fluid, and the exact boundaries of what is and what is not a local network are not always clear. Stallings [2], in a recent text on local networks, makes a contribution to this terminology by identifying, in addition to long-haul networks, the following four categories: multiprocessor systems, high-speed local networks, computerized branch exchanges, and local area networks. Multiprocessors, according to Stallings, are more tightly coupled than are local area networks, and have completely integrated communications and usually some central control. High-speed local networks are usually used in a computer room to provide high throughputs between such devices as mainframe computers. Computerized branch exchanges (CBXs) are digitized telephone networks that can switch data circuits at reasonably rapid rates. Finally, according to Stalling's definitions, *local area networks* have the properties previously attributed to *local networks*.

The majority of this text is devoted to local area networks in Stallings' terminology, although in what follows, the terms *local network* and *local computer network* are often used interchangeably with local area network. Computerized branch exchanges are treated in Chapter 4. Multiprocessor systems and high-speed local networks are not covered.

One further point of terminology needs to be made. Because of the distances involved, some definitions of local networks would seem to exclude broadband networks that operate over a CATV-type system. Such networks are included in local networks for purposes of this text. Along these same lines, it is convenient in discussing random access protocols in Chapter 9 to initiate the discussion by describing protocols normally used on multiaccess radio and satellite links. It would seem that little is lost by including such networks in the category of local networks.

The following list cites several application areas for local networks:

- *Campus computing facilities:* Typical uses include data processing, data interchange, and data storage;
- *Office automation:* Intelligent workstations in this application provide such services as word processing, electronic mail, and electronic copying;
- *Factory automation:* Typical uses include intelligent workstations for engineers, computer aided design, computer aided manufacturing, and inventory control;
- *Library systems:* Typical uses include book inventory and check-out, document retrieval, and electronic copying; and
- *Legal systems:* Typical uses include case retrieval, billing, and word processing.

1.3 Local Network Design and Performance

Physically a local network consists of a communication channel to which a number of intelligent nodes is coupled. The intelligent nodes have several functions: they must transmit and receive signals on the channel; they must control access to the channel; and they must interface the devices coupled to the network at the node. Three types of design problems arise:

1. Design of the network architecture and protocols;
2. Design of the nodal hardware; and
3. Design of the nodal software.

A top-down design begins with step 1. Standards are being developed for a layered architectural structure with three layers for covering the communication aspects of the network. These layers are identified as the *physical layer*, the *media access layer*, and the *logical link control layer*. Briefly, the physical layer deals with generating and transmitting over the channel the physical signals that represent bits. The media access layer manages access to the channel, determining which node can transmit at any given time. The logical link layer operates with bits grouped into packets and manages those functions associated with activities such as error control. More detail on these layers is given in Chapter 10.

Physical layer functions are largely carried out by hardware with some support from software, which usually runs on microprocessor equipment at the nodes. The media access functions tend to use software, again running on microprocessor equipment at the node, with a little support from hardware. The logical link control functions are mostly carried out by software. A complete local network design is a relatively complicated systems problem with subsystems of both hardware and software.

This text is concerned with performance analysis of local computer networks. Performance analysis in this sense has come to mean the development and study of mathematical models that predict the performance of networks in some well-defined sense. As carried out in this text, performance analysis is performed after the fact. That is to say that various common local network designs are taken as given and then analyzed to determine their performance.

As it relates to the design process, performance analysis is carried out in conjunction with the design of network architecture, step 1 in the list of design problems. Usually several designs and their performance are considered and compared. The particular media access technique chosen tends to dictate several different implementations that may have different performance models and thus have different performance.

This book is structured to emphasize performance analysis, which in turn results in detailed studies of media access. An attempt is made, however, to balance this approach in several ways. Chapter 2 is included to cover most of the issues that occur in the designing of the physical layer of local networks. Chapter 10 has a section on logical link control protocols. To carry out performance analysis of the various types of networks, a more or less complete description of each type of network is given.

Thus, although the thrust of the book is performance analysis, some insight into most local network issues is provided.

1.4 Outline of Book Contents

This section describes the contents of each chapter of this book.

Chapter 2 reviews the various aspects of data communication in sufficient depth to make later uses of the material understandable.

Chapter 3 provides the necessary background for modeling the flow of data packets in a network. The chapter covers such material as Little's Law and an introduction to queuing theory.

Chapter 4 is a self-contained treatment of circuit-switched local networks. This chapter presents a quantitative treatment of the problem of PBX sizing.

Chapter 5 is a detailed introduction to local networks. It also discusses assumptions and definitions used in the following chapters and includes brief comments on several practical networks.

Chapter 6 introduces three extreme access protocols that are tractable to analysis and together illustrate the full range of performance.

Chapters 7 through 9 provide detailed discussions of polling, ring, and random access networks. Each chapter begins with a detailed discussion of the design of networks in the class and develops a performance model for networks of the type treated.

Chapter 10 discusses various protocol issues, including the ISO seven-layer model and the IEEE 802 draft standards for local networks.

Chapter 11, the final chapter, contains a case study of random access and token bus networks.

1.5 Summary

This chapter gives a perspective on computer communication networks, discusses the various types, and comments on some of the design problems. Local networks are defined, their characteristics are enumerated, and several application areas are noted.

Design and performance issues for local networks are discussed briefly and the point is made that the text is structured to emphasize performance analysis. Secondary aspects of the book, however, balance this approach with other types of information on local networks. A final section of the chapter outlines the topics covered by the book.