High-Capacity Local and Metropolitan Area Networks

Architecture and Performance Issues

Edited by Guy Pujolle

NATO ASI Series

Series F: Computer and Systems Sciences, Vol. 72

High-Capacity Local and Metropolitan Area Networks

Architecture and Performance Issues

Edited by

Guy Pujolle

Lab. MASI – UPMC-CNRS
University of Paris VI
45, avenue des Etats-Unis, F-78000 Versailles, France

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo
Hong Kong Barcelona Budapest
Published in cooperation with NATO Scientific Affairs Division

Proceedings of the NATO Advanced Research Workshop on Architecture and Performance Issues of High-Capacity Local and Metropolitan Area Networks, held in Sophia Antipolis, France, June 25–27, 1990

ISBN 3-540-53767-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-53767-8 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

45/3140-543210 – Printed on acid-free-paper

High-Capacity Local and Metropolitan Area Networks

Architecture and Performance Issues

NATO ASI Series

Advanced Science Institutes Series

A series presenting the results of activities sponsored by the NATO Science Committee, which aims at the dissemination of advanced scientific and technological knowledge, with a view to strengthening links between scientific communities.

The Series is published by an international board of publishers in conjunction with the NATO Scientific Affairs Division

A Life Sciences
B Physics

Plenum Publishing Corporation
London and New York

C Mathematical and Physical Sciences
D Behavioural and Social Sciences
E Applied Sciences

Kluwer Academic Publishers
Dordrecht, Boston and London

F Computer and Systems Sciences
G Ecological Sciences
H Cell Biology
I Global Environmental Change

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo Hong Kong
Barcelona

NATO-PCO DATABASE

The electronic index to the NATO ASI Series provides full bibliographical references (with keywords and/or abstracts) to more than 30000 contributions from international scientists published in all sections of the NATO ASI Series. Access to the NATO-PCO DATABASE is possible in two ways:

– via online FILE 128 (NATO-PCO DATABASE) hosted by ESRIN, Via Galileo Galilei, I-00044 Frascati, Italy.

– via CD-ROM "NATO-PCO DATABASE" with user-friendly retrieval software in English, French and German (© WTV GmbH and DATAWARE Technologies Inc. 1989).

The CD-ROM can be ordered through any member of the Board of Publishers or through NATO-PCO, Overijse, Belgium.

Series F: Computer and Systems Sciences Vol. 72

Preface

The main objective of this workshop was to review and discuss the state of the art and the latest advances in the area of 1-10 Gbit/s throughput for local and metropolitan area networks.

The first generation of local area networks had throughputs in the range 1-20 Mbit/s. Well-known examples of this first generation networks are the Ethernet and the Token Ring. The second generation of networks allowed throughputs in the range 100-200 Mbit/s. Representatives of this generation are the FDDI double ring and the DQDB (IEEE 802.6) networks. The third generation networks will have throughputs in the range 1-10 Gbit/s.

The rapid development and deployment of fiber optics worldwide, as well as the projected emergence of a market for broadband services, have given rise to the development of broadband ISDN standards. Currently, the Asynchronous Transfer Mode (ATM) appears to be a viable solution to broadband networks. The possibility of all-optical networks in the future is being examined. This would allow the tapping of approximately 50 terahertz or so available in the lightwave range of the frequency spectrum. It is envisaged that using such a high-speed network it will be feasible to distribute high-quality video to the home, to carry out rapid retrieval of radiological and other scientific images, and to enable multi-media conferencing between various parties.

Given this current ferment in the broadband telecommunication networks area, as well as the projection into the future of lightwave networks, it is clear that a number of interesting research issues must be addressed. For instance, protocols must be redesigned and implemented in hardware so that they can be compatible with the high-bandwidth service projected. Another factor that will influence the design of future protocols is the fact that optical communication links have a very low error rate. Invariably, the interest focuses on the end-to-end transport layer of the protocol stack.

Another area that requires extensive research is the way one controls access to the network. In particular, one can visualize a number of different high-bandwith applications utilizing the high-speed network. These applications will have different grade-of-service requirements. For instance, some of them may be tolerant to time delays and intolerant to packet loss, and others may be intolerant to time delays and tolerant to packet loss. In addition, each them may be highly bursty. That is, most of the transmitted information may arrive during short periods of time. Given such a mix of traffic, the question arises of how one controls user access to the high-speed network so as to provide an appropriate service.

The objective of this book is to describe in detail the developments that are taking place. Despite the complexity of the field, this book attempts to cover all its aspects in order to give as complete and accurate a picture as possible.

Guy Pujolle

Table of Contents

DQDB

Performance of ATM Techniques

New Trends in High-Speed Communications

Performance Issues

Gateways

Control in High-Speed Networks

Network Management and Control in Broadband Telecommunication Networks: Research Issues and Some Partial Solutions*

Mischa Schwartz

Department of Electrical Engineering and Center for Telecommunications Research, Columbia University, New York, NY 10027, U.S.A.

Abstract: In this paper we first survey briefly the state of the art in current telecommunication networks. We then focus on ATM-based broadband integrated networks of the near future. Examples include the IEEE 802.6 DQDB metropolitan area network standard and Bellcore's Switched Multimegabit Data Service (SMDS), which uses the 802.6 protocol and ATM at its lowest layers. Looking further into the future it is projected that lightwave- (fiber optic-) based networks may be deployed running at Gbps and Tbps rates. These very high speed networks introduce critical problems of management and control. These include the design of end-to-end lean (lightweight) protocols; access, flow and congestion control; and the real-time detection and identification of network faults when they occur. Some representative ongoing work at Columbia in each of these areas is described, with a hope that it may stimulate additional work elsewhere.

I. Introduction

This paper focuses on research issues in network management and control that become particularly critical in the context of the very high speed, lightwave-based networks of the future under intensive study by many organizations throughout the world. Examples include the design of end-to-end protocols capable of processing packets at the Gbps rates made available to end users in these projected networks; access, flow, and congestion control techniques; and fault management procedures running in real time, in possibly distributed fashion, to automatically detect, identify, and correct problems as they occur.

*Work carried out under the support of National Science Foundation grant CDR 88-11111 and ONR grant N00014-90-J-1289.

NATO ASI Series, Vol. F 72
High-Capacity Local and Metropolitan Area Networks
Edited by G. Pujolle

Note that these networks would be handling high resolution images and video, in addition to data and voice. Provision must thus be made in managing and controlling the networks to take the characteristics of these disparate traffic types into account.

We begin the paper by providing a brief overview of existing digital networks, then move on to the broadband ATM networks proposed for the next decade. We then discuss briefly proposals for very high speed lightwave networks of the future. Both the ATM networks and the much higher bandwidth networks proposed for the future require solutions to the research issues in network management and control noted above. The major portion of the paper summarizes ongoing work of our group at Columbia devoted to research in these and related areas.

II. The Networks to be controlled

Current networks

It is clear that, despite efforts to integrate them, networks worldwide still consist principally of two types: circuit-switched voice networks and packet-switched data networks. Wide-area X.25- based networks are widespread, but the fastest-growing segment of the network market has been that of the 1-16 Mbps LANs, with bridges and routers running at 1.5-2 Mbps rates used to interconnect them. The design and technology is fairly stable by now. Optical transmission is now widespread, but strictly as a replacement for, or enhancement of, cable and microwave traditionally used by the public carriers. Widespread higher-bit rate user communications made possible by the use of fiber is probably still a few years off.

Near-term ATM networks

The rapid development and deployment of fiber optics worldwide, as well as the projected emergence of a market for broadband services, have clearly spurred the recent efforts to develop broadband ISDN standards. ATM, the 125 Mbps FDDI LAN standard, and the 150 Mbps IEEE 802.6 MAN standard are all made possible because of the ability of fiber to potentially provide much higher speed capability in a cost-effective manner. Good-quality video, rapid retrieval of images, scientific visualization, and multi-media conference calls, among many other possibilities, become technologically feasible at the projected ATM rate of 150 Mbps. It is clear from all studies that digital video and moving images have been the driving forces behind the push to these higher bit rates. The actual deployment of ATM networks will require fiber penetrating businesses and homes. This is the remaining stumbling block on the road to broadband ISDN.

The drive toward broadband ISDN and ATM networks in the U.S. comes from two major directions. A number of agencies of the U.S. government have set in motion a plan to interconnect government-supported supercomputer centers using a very high speed fiber backbone. Ultimately, supercomputer users would have high bandwidth connections as well, enabling real-time interaction with supercomputers ("scientific visualization") to become a reality. The other push to broadband ISDN comes from the RBOC's (regional Bell Operating Companies), led by their R & D arm Bellcore. Their strategy is to first focus on the delivery of high bandwidth data services, then to follow with broadband ISDN. To this end Bellcore has, with the help of the RBOC's developed a plan for Switched Multimegabit Data Service (SMDS) [1]. The service, when implemented, would provide a 45 Mbps interconnection capability to LANs, using an IEEE 802.6, ATM-based, interface.

Lightwave networks of the future

A number of telecommunications research organizations worldwide are studying the possibility of all-optical networks in the future that would be capable of tapping the approximately 50 terahertz (THz) or so available in the lightwave range of the frequency spectrum. This represents at least three orders of magnitude more than all of the current radio spectrum, up to and including the microwave band! A variety of fascinating applications come to mind when one visualizes the significance of opening up this relatively untapped reservoir of bandwidth. Past experience indicates that users of such networks, whether at home, business, or institution, will themselves develop undreamed-of applications once presented with this extremely wideband capability.

Our Center for Telecommunications Research at Columbia, in conjunction with a number of our industrial partners, has embarked on a project called ACORN, to develop a prototypical network of the future. Users would be provided with at least 1 Gbps ports (an order of magnitude more than FDDI, 802.6, or the

ATM standard). The network itself would be all-optical, using linear optical elements only (including optical amplifiers to extend the geographical coverage capability). It would provide a packet/circuit switching capability, with any dynamic switching required being done at the network periphery, using high-speed electronics. Our current plan is to deploy a small experimental version of this network in three years time.

III. Network Management and Control: Some Research Issues

Given this current ferment in the broadband telecommunication networks area, as well as the projection into the future of lightwave networks yet to come, it is clear that a number of very interesting research issues arise that must be addressed. We focus here only on those topics relating to network management and control. Activity has already begun on a number of these problems in various organizations.

For example, there is now widespread recognition that protocols will have to be simplified, possibly redesigned, and implemented in hardware to provide compatibility end-to-end with the high bandwidth service projected here. There appears to be a growing realization that the data link layer common in packet-switched networks may become unnecessary in a world of very low error rate optical communication links. Fast, simple, hardware-based routing becomes a requirement in this environment. Hence interest focuses on the end-to-end, transport layer of the multilayer protocol stack. How does one run this layer at the high bit rates discussed here? Investigators have now begun to consider the possibility of lightweight or lean protocols for this purpose. Examples include XTP (The Protocol Engine) [2], [3], NETBLT [4], VMTP [5], and a protocol proposed by investigators at AT&T Bell Laboratories [6]. We describe later in this section our own work on high-speed transport protocols, using a general multiprocessor approach [7].

A second area requiring extensive research is that of access, flow and congestion control. How does one control access to a very high speed network by a number of possibly different bandwidth traffic types, so as to provide the appropriate service requirement for each? ATM access control, on which papers have begun to proliferate, is one example. Here two types of control are usually studied: a static control, to determine, on an average basis, how many calls or services the 150 Mbps broadband ISDN ATM interface to the network, can accommodate; a dynamic control to ensure that the multiplicity of calls, once established, do not overflow the interface. The flow control problem, on the other hand, is the dual to the access problem: how is a particular destination (receiver) to be protected if a multiplicity of sources statistically direct too much traffic at it? Clearly the receiver must flow control the sources to prevent its buffers from overflowing, yet still allow a reasonable flow to come through, as well as provide the proper performance for the various traffic streams. This problem becomes particularly critical at the Gbps and even terabit per second (Tbps) rates of wide area lightwave networks of the future, with literally thousands of packets enroute between a particular source-destination pair. The same problem arises with congestion control.

Take, as a simple example, a 1000-km network running at 1 Gbps. Propagation delay is the order of 5 msec., using a figure of 200,000 km/sec as the effective speed of light over the network. If packets are 1000 bits long (note that ATM cells are 53 bytes or 424 bits long, compounding the problem), there can be as many as 5000 packets enroute over the 1000-km "pipe"! Increasing the bit rate to 10 Gbps raises the number to 50,000! Increasing the packet length helps of course, but the basic problem is still there: how does one control flow over the network in a distributed manner, over relatively long propagation delays in which the ratio of delay to packet length is enormous? We summarize here some preliminary approaches to the problem, developed for a much smaller metropolitan area network [8].

Finally, one last problem we discuss in this paper is that of network management, with particular reference to fault management. This is clearly not solely a problem of high-speed networks, the primary emphasis here. It has, in fact, taken on great significance in the past few years with the proliferation of networks, both public and private. The ISO has been developing a set of network management standards as part of the OSI Reference Model and many vendors have begun to market network management products [9]. The introduction of higher-speed networks can only exacerbate the problem. One must therefore begin to look for ways in which to simplify the process of network management. The work we have begun to carry out, to be summarized briefly here, is that of unifying methods of detecting and identifying faults in networks. The approach we have adopted is that of representing any subsystem, system, or portion of a network under surveillance for possible faults as a finite state machine. A sensor or potential detector of a fault is then itself a reduced state finite state machine. The objective is to detect a known type of fault in real time, when it occurs, as quickly as possible with as small a detector as possible [10].

IV. Network Management and Control: Summary of Work Accomplished

As noted in the previous section, a variety of research issues in network management and control are raised with the advent of high-speed, fiber-based, networks. We summarize in this section some of the work carried out by our group at Columbia in this area. Details appear in the references cited [7], [8], [10].

Very high bit rate end-to-end protocols

Our objective in this work has been to develop a transport protocol implementation that can handle line rates in excess of one Gbps. One approach adopted in the literature has been to design a specialized, simplified protocol specifically designed with hardware implementation in mind [2], [3], [5]. Our approach has been to take a fairly general multi-processing platform and customize it to the needs of a particular transport protocol. In particular, we have focused on providing a multiprocessor capability for the Transport Layer Class 4 (TP4) protocol of the OSI Reference Model 7-layer suite [11].

Multiprocessing can be carried out at various levels: dedicate a processor to each transport connection, assign a processor to specific tasks involved in processing a packet at the transport layer or assign a processor individually to each arriving packet. We have chosen the last possibility as the one with the most potential for truly high speed performance. (Combinations of all three approaches can be adopted as well to further increase the speedup gains attainable).

Given parallel processing at the packet level, a number of options become available from which to choose. These include the scheduling strategy adopted (deterministic or random, for example), as well as the use of shared and/or local memory for holding packets and processing related data structures. We have chosen, for our purposes, to focus on round-robin (deterministic) scheduling with local processor memory. Shared memory is used only to access context records for each transport connection, reducing memory contention considerably.

A high-level view of the receiver portion of the architecture we have designed appears in Fig. 1. (A more detailed description appears in Reference [7]). The packet processors are labeled $P_1, \ldots P_n$. The ILLP, the Input Low Level

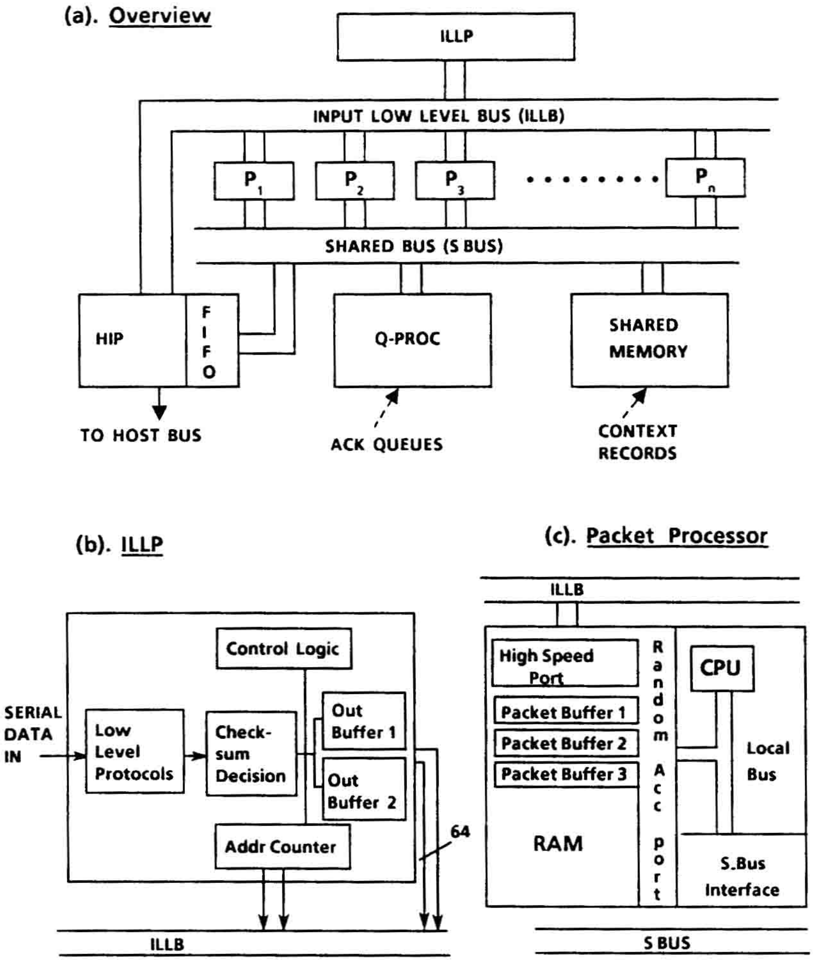

Fig. 1. High-Speed Multiprocessor Architecture

Processor, and an Output Low Level Processor (OLLP) not shown, handle line I/O, error checking, framing, and packet transfer into- and out of memory. If any lower level protocols are required to exist, they are handled by the low level processors as well. The multiprocessor pool handles all transport protocol processing functions. The Host Interface Processor (HIP) shown is a DMA controller to transfer packets to and from host applications. The queue (Q) processor and its associated ack queues are required to handle generation of