# Chapter 1

# Introduction

Due to the advent of fiber optics and the progress in the field of semi-conductors, network technology including transmission speed has been significantly improved. With this improvement, the nature of applications transmitted through networks has also been dramatically changed. These changes include a variety of services such as high speed data transfer, HDTV (High Definition TV), high resolution image communication, video retrieval, video conferencing, multimedia communication, and video on demand, which have various requirements.

The need to deal with this broad spectrum of services has brought about the **B-ISDN** or **broadband integrated services digital network**. B-ISDN is used to support traffic requiring bandwidth ranging from a few Kbit/s (e.g., transmission of slow terminal data) to several hundreds Mbit/s (e.g., transmission of moving image data) and is also required to meet the diverse service and performance requirements of multimedia traffic. For instance, real time voice requires rapid transfer, but the loss of small amounts of voice information is somewhat tolerable. In some applications, real time delivery is not of primary concern, but high throughput and strict error control are required. Other services, such as real time video communications, require error tolerant transmission as well as rapid transfer. B-ISDN is an efficient network standard to support and to deal with all of those services in the same way, which should eliminate the need for separate networks for the different services and, instead, integrate them into one universal network.

In order to satisfy the requirements of these applications, **ATM (asynchronous transfer mode)** has been accepted as the switching and information transmission technique for implementing the B-ISDN. ATM is a packet-based scheme which deals with a fixed-sized packet, called a "cell". In a packet-switched network, network resources are used more beneficially than in a circuit-switched network because the network resources are not monopolized by a single user. Moreover, ATM exploits network resources much more efficiently by means of statistical multiplexing.

However, unlike other packet-switched networks, ATM provides guaranteed network resources in order to meet the requirements of quality of service negotiated at the beginning of a session. Because of its features, ATM is one of the most prominent and efficient transmission techniques to support current backbone networks.

Meanwhile, from the need to interconnect many disparate physical networks and to make them function as a single virtual network, internetwork (or internet) technology first emerged about thirty years ago. Ever since, with the development of internetworking protocols like IP (internet protocol) and ISO's Connectionless Network Protocol (CLNP), the TCP/IP or Transmission Control Protocol/Internet Protocol (commonly referred to as the Internet) has made it possible to communicate across any set of interconnected networks all around the world. Since its development in the late 1970s, the TCP/IP protocols have achieved substantial growth based on the open systems interconnection promise of OSI, so that it is now one of the the most predominant internetworking protocol for LANs [COME88][HAND94][KESS93].

In order to provide reliable services over the Internet, flow control schemes in the TCP layer have attracted much attention recently. Hence, a plenty of studies based on TCP's control schemes have been conducted, too. On the other hand, when the TCP packets are serviced by an ATM network, the packets from the TCP layer have to be fragmented appropriately into a series of ATM cells, because, hierarchically TCP/IP is located in a higher layer than the ATM layer as shown below.

| |
|---|
| TCP (Transmission Control Protocol) |
| IP (Internet Protocol) |
| ATM (Asynchronous Transfer Mode) layer |

While TCP supports reliable services by means of packet-by-packet acknowledgment, ATM does not support end-to-end cell-by-cell acknowledgment. Thus, because of the fragmentation, if a cell is dropped in between the links along the route, the transmission of the corresponding TCP packet results in useless transmission. In order to avoid this fragmentation problem, some types of flow control schemes in ATM have been studied by various researchers. The solution these control schemes employ is to drop an entire packet before it reaches destination if the network is deemed to be congested and the packet is to be partially dropped. These schemes, however, delete the packet by a simple decision. Thus, they possibly discard more packets than necessary and decrease utilization of network resources.

This dissertation first describes the generalization of the existing congestion control schemes. In the process of this generalization, this author presents a new optimistic congestion control scheme named the "probabilistic delayed packet discard" (PDPD) scheme. The PDPD scheme delays the actual packet discard operation based on probability. This scheme should achieve higher throughput than the existing ordinary discard schemes by setting the probability adaptively according to the network status.

Chapters 2 and 3 describe the background of this study. In Chapter 4, this author discusses the generalization of packet discard schemes and introduces the "Probabilistic Delayed Packet Discard" (PDPD) scheme. The simulation results are analyzed in Chapter 4 to show the throughput improvement with the PDPD scheme. Chapter 5 is the conclusion.