CCIT Report #407 December 2002

Abstract

In this work, we address the problem of reliable multicast distribution of bulk data to a large set of independent clients. The clients are not coordinated; i.e., they may start data reception at any point in time. Clients are heterogeneous in the sense that their maximum achievable data rates from the server vary widely, both among clients and with time. Moreover, the network may become congested and packets may be lost. Our goal is to permit each client to receive all the data as early as possible, while minimizing network resource utilization. The problem being addressed represents mass distribution of bulk data, such as software, software updates, content distribution, etc. This work demonstrates that any simple, erasure correcting codes computed over small blocks of data, can be used for this purpose in a near optimal way.

Our transmission scheme comprises a set of co-scheduled channels, whose transmission rates increase exponentially. Every channel carries all the data cyclically, along with redundant packets computed using an erasure correcting code. (Note that, due to the perpetual nature of the transmission, the use of redundancy does not cause the transmission of any "extra" data.) This code is applied to small, fixed-size groups of packets. The key element in this scheme is the partitioning into small groups, which permits the efficient use of standard codes, the smart scheduling of the transmission within each channel, and the interleaving of the channels. We show that a user that subscribes to any contiguous set of channels including the slowest one must only receive an amount of data equal to the original file size, regardless of starting time. This remains nearly true even if packets are lost or the user changes its subscription. At the same time, the network is used efficiently, in that the data rate carried over a given link is no greater than that required by the fastest downstream subscriber. The basic scheme is then extended in several ways to permit finer granularity of the user data rates, and is shown to remain near-optimal even if network conditions and subscriptions change dynamically. One of these extensions also applies to schemes that use an erasure correcting code that is computed over the entire file. Finally, the scheme is highly scalable because the only interaction with users is their subscription to channels.