CS 336

Assignment 3: Broadcast Local Area Networks

Due the week of September 26/27, 2012

This week, we will explore the operation of the most widely used local area network technology, Ethernet, and begin to examine wireless networking technology.

I would like you to read §2.6 of Peterson and Davie this week. The first section covers Ethernet. The treatment of Ethernet is rather brief because they present the algorithms that are used in Ethernet without exploring the performance of these algorithms. The additional readings described below are designed to help you explore these performance issues.

The first is another "Classic" paper, "Ethernet: Distributed Packet Switching for Local Computer Networks" by R.M. Metcalfe and D.R. Boggs. This is the paper describing the original experimental system out of which "Ethernet" emerged. Those who want a bit more background on the invention of the Ethernet might enjoy "Dealers of Lightning: Xerox PARC and the dawn of the computer age" by Michael Hiltzik, a history of the research center at which Ethernet and many other computer technologies we take for granted today were first developed. This book is available in Schow. (It is not a required reading.)

The second paper has the distinction of having one of Ethernet's inventors as a co-author. The paper is "Measured Capacity of an Ethernet: Myths and Reality" by D.R. Boggs, J.C. Mogul and C.A. Kent.

The performance of an Ethernet depends on when the computers connected attempt to transmit data and on how much data they attempt to transmit. These details cannot be predicted exactly. Instead, analyses of Ethernet performance attempt to predict the network's expected behavior based on probabilistic models of the behavior of the attached computers. As a result, the appendix from Walrand's computer networks textbook that I provided as an optional reading last week may again be helpful.

Exercises

1. The following question is taken from Tanenbaum's text. He describes a scenario involving multiple frames, but (for reasons suggested by the next problem) I would like you to limit your attention to the first pair of frames sent and to assume that where Tanenbaum says "mean" he meant to say "expected."

Two CSMA/CD stations are trying to transmit long (multiframe) files. After each frame is sent, they contend for the channel, using the binary exponential backoff algorithm. What is the probability that the contention ends on round k, and what is the mean number of rounds per contention period?

In addition, please provide a formula for the mean number of "slots" per contention period.

Warning: The goal of this question is to get you to see how complex the behavior of the exponential backoff algorithm really is. My hope is that this will make you appreciate the need for some of the approximations in the analysis of Ethernet presented in the original paper. As a result, the answers to this question are not nice and simple. Expect nasty products or summations. Don't spend a lot of time trying to simplify terms, just try to mathematically capture the behavior of a real Ethernet as closely as you can.

- 2. These is an interesting behavior exhibited by the binary exponential backoff algorithm known as the Ethernet capture effect.
 - (a) Complete parts a, b, and d of question 2.44 from Peterson and Davie.
 - (b) Obviously, I don't like part c of question 2.44. It is just a bit too vague and/or difficult mathematically. As an alternative, I would like you to try to take part b a bit further. For parts a and b, you investigated the probability that A would win a round *immediately* after the first collision of a "backoff race". You should have found that these probabilities are fairly high. If you drop the word "immediately", however, the probability is even higher because if during a given backoff race A and B collide on both their first and second attempts, A will still have a higher probability of winning on the third, fourth, and subsequent attempts than B.

What I would like you to do is find a reasonable lower bound for the probability that A will ultimately win the second backoff race described in part b, either immediately or after several attempts. (Hint: I think the best way to approach this is to find an upper bound on the probability that B might eventually win and use 1 minus this upper bound as your lower bound on A.)

- (c) In their discussion of Ethernet performance, Boggs, Mogul and Kent make a statement that suggests the scenario suggested by the Peterson and Davie question could not occur in their environment. Explain why. Could a similar situation occur in their environment? Explain.
- 3. While the layers of the OSI model are designed to avoid tackling all the issues involved in building a network at once, it is sometimes helpful to look at how the layers interact (or should interact). In their paper, Boggs, Mogul and Kent give a hint of the importance of one such interaction between the medium access protocol used and the retransmission scheme used. This occurs in the last paragraph of section 4.3.

Explain precisely how the problem they describe could occur. How difficult would it be to incorporate their suggestions in the ARQ protocols we studied last week.

4. Boggs, Mogul, and Kent clearly state that the failure to understand the correct interpretation of theoretical analyses of Ethernet behavior has led to "myths", but they never quite clearly state whether their results directly conflict with any published analytical results.

One of them does.

Consider the data shown in the very first graph presented in their paper, Figure 3.3. It shows that the efficiency achieved with 24 stations transmitting 64 byte packets is 85%. The analysis in the Metcalfe and Boggs paper provides the means to predict this efficiency. If, as they assume in their analysis, binary exponential backoff effectively estimates Q, then the analysis given in §6 of Metcalfe and Boggs should enable you to estimate the efficiency for the situation where 24 stations are sending 64 byte packets.

- (a) Use the results in §6 of Metcalfe and Boggs to calculate an estimate of the efficiency for the situation where 24 stations are sending 64 byte packets. Your estimate should fall far below 85%.
- (b) One possible explanation for the mismatch between your calculations and the measured efficiency is the Ethernet capture effect explored in two of the earlier problems in this assignment. In their paper, Metcalfe and Boggs assume that the "Ether's time is divided between transmission intervals and contention intervals." The formula they give for efficiency assume

the network alternates between transmissions and contention. If the capture effect is significant, this assumption would be invalid. Instead, the network would alternate between periods where a station that had captured the network transmitted several packets, winning each contention round between packets in a single slot, and periods of true, fair contention. Propose a formula for the efficiency of an Ethernet in this case. You will have to include an extra variable, S, denoting the expected number of consecutive transmissions made by a station while it has captured the network. Remember, that in their analysis, Metcalfe and Boggs made significant assumptions/approximations. You should not be surprised if you need to do the same. You should, however, state such assumptions clearly.

- (c) Using your formula and the data for the situation where 24 stations are sending 64 byte packets, estimate the value of S.
- 5. Warning. This is a very open-ended question.

In their paper, Boggs, Mogul and Kent question the validity of the results published by many others. Anyone who does this certainly deserves to have their own work subjected to similar scrutiny. I'd like you to take on the job.

I'd like you to imagine what you might do if you were one of them getting ready to submit the paper for publication. The paper depends on a great deal of data collected by programs that were probably complex enough to contain a bug or two. What if the bugs affected the data? This is a risk one must address when using software to collect data in an experiment or to simulate a system to predict its behavior.

There are many ways to validate such data. One approach is to look for properties of the data that can be predicted from "theory" and see if the data collected agrees with these prediction. Normally, it is not possible to predict all the data from theory. One can, however, often find limited cases or relationships between data collected in different scenarios that can be predicted.,

For example, while the general behavior of an Ethernet may be too difficult to precisely analyze mathematically, one might be able to say something precise about situation where only two stations are active. Alternately, while one might not be able to predict the exact utilization with 25 stations, one might be able to predict the ratio of the utilization with 25 stations sending 1k packets to that with 25 stations sending 2k packets. If one makes such a prediction and the data contradicts it, then it is time to question the accuracy of the experiment.

For this problem I am asking that you try to find a few ways in which you can either argue mathematically that their data meets expectations or that there appear to be surprises. To some extent, problem 3 is an example of what I had in mind (but don't use that example for this question). If you find any surprises, did the authors notice them as well? Do they explain them in the paper?