CS640 Project Proposal

Aubrey Barnard Perry Kivolowitz

February 26, 2009

1 Introduction

For our class project in CS640, Computer Networks, we propose to investigate strategies for scheduling outgoing network traffic among multiple interfaces on a single client computer. Scheduling is an integral part of making multiple network interfaces simultaneously available to clients, greatly improving their networking capabilities and performance.

1.1 Problem

Increasingly, personal computers have multiple network interfaces available for use. Unfortunately, current operating systems, and hence applications, do not use more than one of these interfaces at once to improve network performance.[1] Finding a way to eliminate this under-utilization of network resources would bring great benefit, but involves solving several problems. We hope to take the first step toward better utilization by addressing the problem of scheduling traffic among multiple, heterogeneous interfaces.

This problem factors into two smaller sub-problems. The first sub-problem is how to characterize the performance of a network interface, and how to track changes to its characteristics over time. This involves developing and evaluating metrics for measuring various properties of network links. The second sub-problem is how to best schedule network traffic given the varying characteristics of the multiple interfaces. This involves analysing network measurements and formulating strategies for efficient scheduling given the measured network conditions.

[1] Performance can be measured in many ways, such as bandwidth or reliability. We hope our work will be able to address multiple performance dimensions.

Our end goal is to develop an algorithm for scheduling network traffic among multiple network interfaces given each interface's (changing) characteristics. We will know our algorithm works if it allows a computer using multiple network interfaces to perform better than a computer using a single interface, for some performance measure.

1.2 Significance

This problem is important because its solution helps pave the way for two significant advances in networking technology: (1) improving network performance by taking advantage of multiple interfaces and (2) improving client mobility by distributing and transitioning TCP connections over multiple interfaces.

2 Approach

There are numerous technical challenges that make it impractical for us to directly distribute traffic to multiple interfaces. Therefore, we plan to simulate the situation with a client and server under our control. This will allow us to accurately measure network conditions and even introduce failure modes. We realize that the realism of our data depends on the realism of our simulation and are confident that our simulation will accurately model reality.

In our simulation, the client will perform four functions. First, it will sample network activity on the client computer. Sampling will record the salient features of the packets but will not store the payload. Second, the client will replay captured traffic to the server. Specific features of the captured outgoing traffic will be duplicated such as packet timing,

size and type. Third, the client will prepare a catalog of what it has just replayed and send this catalog to the server. Finally, the client will listen for a characterization of the traffic the server received.

The server will perform five functions. First, it will capture traffic directed to it by various instances of the client application. Second, it will build a catalog of the captured traffic. Third, it will receive the client's version of the catalog. Fourth, it will compare the two catalogs to produce a characterization of the network environment. Finally, it will send this characterization back to the client, whereupon the client can decide how to better allocate its outgoing resources. Such a cycle of measurement, analysis, and scheduling decisions is the essence of our simulation.

We envision taking a few approaches in developing scheduling strategies. All approaches will "examine" the characterizations of the network links and then make scheduling decisions. One approach will be to develop a "hard" scheduling algorithm. We may also develop a "soft," adaptive algorithm that can adjust its behavior to better meet scheduling needs. Another approach is applying machine learning techniques to the characterization data and learning a classifier that will classify (schedule) traffic.

We have chosen to use the C# programming language for implementing the client and server. It provides ease of development, managed resources, and extensive programming libraries, which should ease and encourage our development. The choice of C# dictates the choice of the Windows platform. Low level capture and retransmission will be performed with SharpPCap which is a C# wrapper to WinPcap, a Windows version of well-known libpcap for Unix.

3 Milestones and Schedule

We identified the major tasks of our project and developed a schedule to help structure our work. Below, we list the tasks along with the dates planned for their completion.

February 26 Proposal submission

March 05 Proposal evolution; definition of metrics; base client and server implementation

March 19 Packet capture

April 02 Client and server running and taking measurements

April 16 Data collection, analysis, characterization; strategies for network utilization

April 30 Strategy evaluation and evolution; machine learning integration

May 10 – May 13 Presentation

May 15 Paper submission

We consider the minimal requirements of our project to be developing metrics and using them to evaluate and characterize network traffic. We would like to achieve the more successful result of analyzing traffic patterns, formulating strategies for allocating network traffic to multiple network interfaces, and developing algorithms to implement those strategies. We would also like to explore a parallel approach using machine learning, but consider this more an extension than a core requirement. These three levels provide us with points to fall back to in case we encounter difficulties.

4 Deliverables

We anticipate creating three deliverables.

Presentation We will present our research to the class. We will negotiate the scope and format of the presentation with Professor Banerjee. In any case, we aim to provide a thorough overview of our work as well as an intuitive explanation of the more technical details.

Paper We will document our research and important results in a paper. We hope to include metrics for characterizing network links and algorithms for effectively utilizing multiple network interfaces.

Source code We will provide the source code and related mechanisms that we used to accomplish our research.