Building a High-Quality Mobile Telemedicine System using Network Striping over Dissimilar Wireless Wide Area Networks

Asfandyar Qureshi, Ali Shoeb and John Guttag
MIT Computer Science and AI Laboratory
{asfand, ashoeb, guttag}@csail.mit.edu

Abstract— **This paper describes an approach to building a high-quality mobile telemedicine system that overcomes the limitations of individual public wireless data networks. Public wireless data channels do not have the capacity to handle the high-bandwidth video needed for applications such as remote evaluation of trauma and stroke patients. Network striping allows us to aggregate multiple physical channels to meet the bandwidth requirements for the video. We have developed flexible network-striping software middleware, and are building a telemedicine system using that middleware. Our approach uses existing communications infrastructure and conventional-off-the-shelf components, making the system easy to deploy.**

I. Introduction

When appropriate health professionals cannot be physically present to diagnose patients, telecommunications technology can be used to connect these professionals to those in need of their expertise. In hospitals and homes, telemedicine has been shown to reduce the cost of health-care and increase efficiency through better management of chronic diseases, shared health professional staffing, reduced travel times, and fewer or shorter hospital stays [1].

We focus on using telemedicine to improve outcomes when dealing with acute incidents involving critical care patients. By relaying real-time telemetry and video from ambulances, we hope to provide EMS teams with expert opinions on complex trauma injuries, and to aid the in-hospital teams in better preparing themselves for incoming patients. In many situations, the timely application of an appropriate therapy is of critical importance, and mobile telemedicine can help.

Our overall research project aims to provide advanced remote diagnostic capabilities for patients in an ambulance moving about in an urban area. From the moving ambulance, we want to transmit real-time uni-directional video, bi-directional audio, and multiple uni-directional physiological streams (EKG, blood pressure, etc). Not only must we stream high-bandwidth video over a wireless channel, we need to handle the additional complications introduced by vehicular motion.

In order for the telemedicine system we design to be useful, the system must be economically viable to build, deploy and operate. We therefore use existing wireless communications infrastructure. We also build our system out of conventional off-the-shelf components.

In most urban areas, a number of public carrier cellular Wireless Wide Area Networks (WWANs) provide mobile connectivity to the Internet. Unfortunately, present-day WWAN channels have low bandwidths and provide little in the way of network Quality-of-Service (QoS) guarantees. Additionally, different WWAN channels can provide very different network QoS. Network QoS is affected by the layout of a provider's wireless access-points, and by the WWAN communication technology, both of which may vary across providers. For these reasons, video-conferencing using a single WWAN channel results in a video quality that is inadequate for many medical purposes.

Network striping, allows the construction of a high-bandwidth virtual channel from a collection of low-bandwidth physical channels. Although much past research has dealt with network striping, many assumptions made by those researchers do not hold when striping over a set of dissimilar WWAN channels. We have therefore developed novel middleware to allow applications to flexibly stripe data over WWAN channels [3], [4]. We are using this to build our telemedicine system.

The available diversity in present-day WWAN environments makes network striping an especially appealing approach. By taking advantage of service provider diversity, overlapping coverage, and network technology diversity (e.g., GPRS and CDMA), one can attempt to provide each application with the illusion that a reliable stable high-bandwidth channel is available. The existence of technological and provider diversity—a benevolent side-effect of the competitive nature of the cellular provider market—is likely to bolster the virtual channel, making it more reliable. The underlying channels are more independent than if the same technology or the same provider were being used.

The remainder of this paper is structured as follows: we outline a motivating application; provide an overview of the nature of WWAN channels; discuss how our end-to-end design eases deployment and provides flexibility; describe important services provided by our striping middleware; and finally document some experimental results, before concluding.

II. Motivating Application

Our overall research project aims to use tele-presence to provide advanced remote diagnostic capabilities for patients in ambulances. This has the potential to improve emergency medicine.

Emergency and trauma medicine focus on situations where urgent or immediate care is required. While not usually providing long-term care, emergency medicine physicians and pre-hospital personnel still provide care with the aim of improving long-term patient outcome.

Emergency and trauma medicine are inherently multi-disciplinary and time critical. Unfortunately, involvement of multiple disciplines at multiple sites encourages breakdowns in communication and continuity of care. Communication between the EMS pre-hospital team and the in-hospital trauma team is especially challenging. Today, audio is often the only available communication medium. A system that relayed real-time telemetry and provided a high quality video link to the two teams could lead to improved patient care by:

1) Providing the in-hospital trauma team with a better advanced understanding of the patient's condition, thereby enabling them to be better prepared to provide effective interventions. They could, for example, avoid unnecessary delays by by-passing the Emergency Department and taking critically injured patients directly to the operating room, ICU, CT scanner, etc.
2) Providing the EMS team with expert opinions on complex trauma injuries from in-hospital physicians, who will be able to observe the situation in real time, and recommend interventions to the EMS team.
3) Reducing the time needed in handing off the patient from the EMS team to the in-hospital team.

Experience with inter-medical-center and mobile remote medicine suggests that real-time video distributed over a high speed Internet connections can be a valuable aid in diagnosing and improving treatment. For example, in the case of suspected stroke victims [5], [2].

III. Wireless Wide Area Networks

In most urban areas, there are a large number of public carrier wireless channels providing mobile connectivity to the Internet. These providers have overlapping coverage areas, allowing us to connect to more than one provider at the same time.

Individual cellular WWAN channels do not provide enough bandwidth for our telemedicine application. The upstream bandwidth[1] offered by these channels is rather limited, and usually less than the downstream bandwidth. For example, the best WWAN channels available in the greater-Boston area provide a little over a hundred kilobits per second of *peak* upstream bandwidth. Our experiments [4] show that achievable average bandwidth is normally lower and varies significantly over time.

We do not expect WWANs to improve enough in the near future to meet the upstream bandwidth requirement of our telemedicine application. Our application's bandwidth demand is fairly extreme for WWANs: our goal is to be able to deliver video whose quality is close to broadcast television quality.

Co-located WWAN interfaces can each provide very different network QoS. US providers use standard technologies, such as `GSM/GPRS` and `CDMA2000`. Our experiments [4] in Boston show that `CDMA2000` channels are far better than `GPRS` channels. A `CDMA2000` link has six times the upload bandwidth of a `GPRS` link.

Packet latencies are important because high packet latencies limit the interactivity of a video-link. For small packets, we have measured minimum packet round-trip-times of around $315ms$ on the `CDMA2000` link and round-trip-times of $550ms$ on the `GPRS` link. For larger 768-byte UDP packets, on both types of channels we have measured round-trip-times of around $800ms$ with a standard deviation of around $100ms$.

Vehicular motion significantly impacts the QoS on both kinds of links [4]. Motion causes average throughputs on both types of links to decrease and become more variable. Motion also affects packet latencies on the `GPRS` links, causing average latencies to rise and become more variable.

IV. Network Striping

Since individual public wireless data channels are unable to handle the demand of our telemedicine application, we aggregate several of these physical channels to construct a high-bandwidth virtual channel.

Inverse multiplexing, or network striping, allows the construction of a high-bandwidth virtual channel from a

[1]Upstream bandwidth refers to how much data one can *send* from a mobile node.

collection of multiple low-bandwidth network channels. Network striping takes data from the larger source virtual channel and sends it in some order over the smaller physical channels.

Much work has been done on network striping [6], [7], [8], [9], [10], [11]. Most of this work focuses on aggregating bandwidth and assumes that the underlying links are relatively stable and similar. Our research in network striping is unique in that we tackle the problems of striping over a heterogeneous set of unstable WWAN channels, and in that we try to accommodate application-specific network QoS sensitivities.

To build our telemedicine application, we have developed *Horde*, software middleware that facilitates flexible network striping over unstable dissimilar channels.

A mobile telemedicine system based on Horde would be fairly easy to deploy. No special communications infrastructure need be set up; all that would be required is a computer on the ambulance with off-the-shelf network interfaces, equipped with standard cellular data plans. Our software is designed to deal with a diverse set of network interfaces, so the virtual channel's bandwidth would scale up as the cellular networks improve.

Additionally, using a set of channels instead of a single channel provides an important choice to the maintainers of the telemedicine system: they can decide how much bandwidth they want, and provision accordingly, without having to involve the carriers in this decision. If a particular ambulance decides to use a roughly 200 kilobit-per-second connection for the video[2], they can achieve this by acquiring two standard data service plans from different `CDMA2000` providers; if instead, they decide to implement a higher-quality video link, they only need to acquire additional plans and the new interfaces into the ambulance's on-board computer. Because of the end-to-end design of our system, no negotiations with the carriers are needed for this bandwidth flexibility.

Horde provides many services to applications.

A. High Bandwidth and Reliability

The primary service provided by Horde is that of a high-bandwidth reliable virtual channel. By taking advantage of service provider diversity, overlapping coverage, and network technology diversity (e.g. GPRS and CDMA), Horde attempts to provide each application with the illusion that a reliable stable high-bandwidth channel is available. Striping allows applications to sustain high throughputs even as individual channel connectivity varies. When a WWAN channel suffers a disconnection (a dropped call), the virtual channel's bandwidth is reduced, but the application can continue.

[2]A video codec like the Windows Media 9 series codec [12] can compress broadcast-quality 1280x720 pixel/frame 25 fps video down to around 8 Mbit/s.

B. Modulating Network QoS

Horde allows applications to describe QoS based policy objectives that the striping mechanism attempts to satisfy. Horde can be used by a set of application data streams, each with its own QoS policy, to flexibly stripe data over a heterogeneous set of dynamically varying network channels. With such channels, different striping policies can provide very different network QoS to different application data streams [3], [4].

In our telemedicine system, we want to transmit real-time uni-directional video, bi-directional audio and uni-directional physiological data streams (EKG, blood pressure, etc). Some of these data streams are less sensitive to network latency (e.g. the EKG data) and others are more sensitive (e.g. the video and audio streams). This implies that the video stream should be preferentially sent over low-latency channels. Furthermore, an encoded video stream may contain both *reference* frames and *delta* frames. In order for delta frames to be decoded properly at the receiver, the frames they depend on must have been successfully transmitted across the network. An application derives more *utility* from a reference frame than from a delta frame. This suggests that reference frames should be sent on more reliable channels.

The key technical challenge in Horde is giving applications control over certain aspects of the data striping operation (e.g., urgent data should sent over low latency channels or critical data over low loss channels) while at the same time shielding the application from low-level details. Horde aims to decrease the programming costs associated with building complex mobile systems that use network striping.

V. Experimental Evaluation

We have implemented a preliminary version of the Horde middleware, and are working on building a mobile video streaming system using this implementation. In this section we report on results from a single experiment in which Horde was used to send data to a host on the MIT network.

The experiment was conducted using a Linux laptop computer connected to a single `CDMA2000 1xRTT` interface and multiple `GSM/GPRS` interfaces. All `GPRS` interfaces were connected to the same service provider; the `CDMA` interface was connected to a different provider. The `GPRS` interfaces used standard cell-phones connected to the laptop over a bluetooth link. The `CDMA` interface was a PCMCIA-based modem. The devices

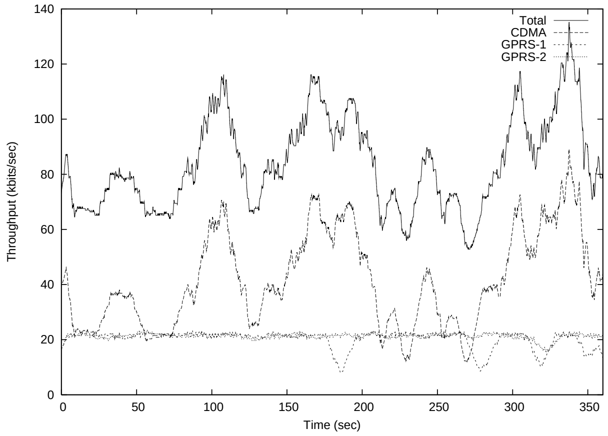

Fig. 1. Throughput provided by Horde in a stationary experiment, using three colocated interfaces.

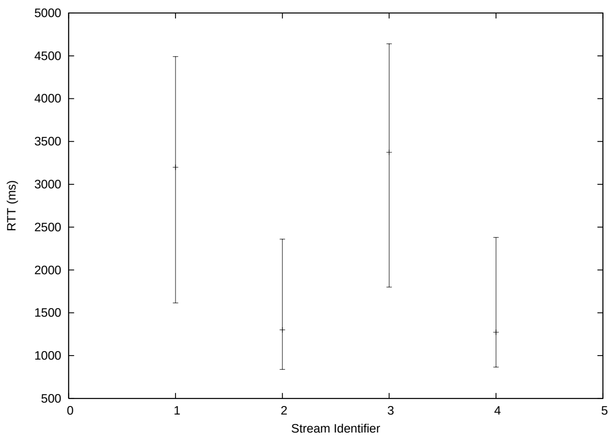

Fig. 2. Round-trip-time distributions for different data streams. The graphs show the median and the upper and lower quartiles for the network round-trip-time experienced by packets in each stream.

were in close proximity to one another and did not use specialized antennas.

Figure 1 depicts the aggregate throughput provided by Horde, and the throughputs provided by each individual WWAN channel in the experiment. The throughputs were calculated using a 10-sec sliding window average.

Figure 2 shows how Horde can be used to modulate stream QoS. Four data streams, each with the same throughput, were created in the experiment. Two of these streams (2 and 4) were tagged as being latency sensitive. Conventional striping systems would have provided all four streams with similar packet latency distributions. In contrast, as figure 2 shows, Horde was able to transmit the latency sensitive streams' packets in lower latency transmission slots.

VI. Conclusions

The deployment of a high-bandwidth low-latency mobile telemedicine system has the potential to reduce medical complications suffered by patients outside of medical environments. Individual available wireless data channels do not have the capacity required for high-bandwidth video streaming. We have, therefore, developed a flexible network striping middleware that aggregates these dissimilar low-bandwidth physical channels, providing a robust high-bandwidth virtual channel to applications. We are in the process of building the telemedicine application on top of this middleware. The end-to-end nature of our system makes it easy to deploy and operate.

VII. Acknowledgments

The authors would like to thank the National Library of Medicine (NLM) and the Center for Integration of Medicine and Innovative Technology (CIMIT) for funding. We would also like to thank George Velmahos, M.D., for providing invaluable insight into how our technology might improve trauma care. Special thanks to Michel Goraczko and Dorothy Curtis.

References

[1] "American Telemedicine Association." http://www.atmeda.org/.

[2] M. P. LaMonte, Y. Xiao, P. Hu, D. Gagliano, M. N. Bahmouth, R. D. Gunawardane, C. F. Mackenzie, W. R. Gaasch, and J. Cullen, ""Shortening time to stroke treatment using ambulance telemedicine: TeleBAT"," *Journal of Stroke*, 2004.

[3] A. Qureshi and J. Guttag, "Horde: Separating Network Striping Policy from Mechanism," in *ACM MOBISYS*, 2005.

[4] A. Qureshi, "Flexible Application Driven Network Striping over Wireless Wide Area Networks." MEng Thesis, Massachusetts Institute of Technology, March 2005.

[5] Y. Xiao, D. Gagliano, M. P. LaMonte, P. Hu, W. Gaasch, R. Gunawadane, and C. Mackenzie, ""Design and Evaluation of a Real-Time Mobile Telemedicine System for Ambulance Transport"," *Journal of High Speed Networks*, pp. 47–56, September 2000. http://hrfp.umm.edu/Papers/2000/mobiletelemed.pdf.

[6] P. Fredette, ""The Past, Present, and Future of Inverse Multiplexing"," *IEEE Communications Magazine*, pp. 42–46, April 1994.

[7] J. Duncanson, ""Inverse Multiplexing"," *IEEE Communications Magazine*, pp. 34–41, April 1994.

[8] H. Adiseshu, G. M. Parulkar, and G. Varghese, ""A Reliable and Scalable Striping Protocol"," in *SIGCOMM*, pp. 131–141, 1996. http://citeseer.nj.nec.com/adiseshu96reliable.html.

[9] A. Snoeren, ""Adaptive Inverse Multiplexing for Wide-Area Wireless Networks"," in *IEEE conference on Global Communications*, pp. 1665–1672, 1999. http://citeseer.nj.nec.com/snoeren99adaptive.html.

[10] P. Rodriguez, R. Chakravorty, J. Chesterfield, I. Pratt, and S. Banerjee, "MAR: A Commuter Router Infrastructure for the Mobile Internet," in *Mobisys*, 2004.

[11] L. Magalhaes and R. Kravets, ""Transport Level Mechanisms for Bandwidth Aggregation on Mobile Hosts"," in *ICNP*, 2001.

[12] "Windows Media 9 Series Audio and Video Codecs." http://www.microsoft.com/windows/windowsmedia/9series/codecs.aspx.