Turning the Postal System into a
Generic Digital Communication Mechanism

Randolph Wang, Kai Li
Computer Science Department
Margaret Martonosi
Electrical Engineering Department
School of Engineering and Applied Science
Princeton University
Princeton, NJ 08544

January 8, 2004

A Project Summary

Recent studies [9, 6] find that as the adoption rates of the Internet and broadband connections slow down in the U.S., the "digital divide" could be solidifying: those with modest incomes, rural residents, and minorities are among those who lag behind in Internet access. Most people living in developing regions, who represent an overwhelming majority of the world's population, also largely fall on the "wrong" side of the digital divide and are missing out on many important opportunities that the digital world has to offer.

Bridging this digital divide, especially by attempting to increase the accessibility of broadband connectivity, can be very challenging. The improvement of wide-area Internet bandwidth is constrained by factors such as how quickly we can dig ditches to bury fibers in the ground. The cost of furnishing "last-mile" wiring can be prohibitively high, and the progress has been excruciatingly slow.

In this proposal, we explore the use of digital storage media (such as DVDs or even hard disks) transported by the postal system as a general digital communication mechanism. While the idea of sending digital content via the postal system is not a new idea – companies (such as AOL.com and netflix.com) have used this approach to deliver software and movies on a large scale for some time – none of these existing attempts have turned storage devices delivered by the postal system into a *generic* communication channel that can cater to a wide array of applications. We shall call such a system a *Postmanet*.

Compared to traditional wide-area connectivity options, the Postmanet has several important advantages. These include wider reach, greater bandwidth potential, low cost, and better scalability. In terms of reach, the postal system is a truly global "network" that reaches a far greater percentage of the world's human population. In terms of its bandwidth, the Postmanet advantage stems from fundamental technology trends: the explosive growth of storage technology density implies that the amount of information that can fit into a fixed amount of volume, or the amount of information that can be shipped by the postal system for a fixed cost, increases exponentially at roughly the same rate, a rate that the wide-area Internet bandwidth growth is unlikely to be able to keep up with. Indeed, far from being a temporary fluke, the bandwidth gap between the Internet and the Postmanet is likely to widen as the storage density continues its rapid improvement.

We would like to use the Postmanet to extend and complement the existing Internet to support a wide array of bandwidth-intensive applications. These may include email with large attachments, web pointing to or embedded with large data objects, remote file system mirroring for sharing and/or backup, peer-to-peer file sharing, video "almost on-demand," publish/subscribe systems for other types of content, and distance learning.

At a first glance, it may appear deceptively simple for an end user to manually burn and ship DVDs for whatever applications he desires. But with many applications, hundreds of communicating parties, and thousands of "messages," for example, manual management would become infeasible. Therefore, systematic transparency support at the levels of network transport, programming models, routing, and applications is necessary. We shall see two recurring themes at these different levels of the system. One is the simultaneous exploitation of the low-latency low-bandwidth Internet and the high-latency high-bandwidth Postmanet so we can combine their latency and bandwidth advantages. The other is the exploitation of the abundant capacity and bandwidth of the Postmanet to improve its latency, reliability, and cost.

We address the NSF *intellectual merit and broader impacts* more fully in Section C.7. We highlight several issues here. (1) An important barrier to broader adoption and use of information technology across geographic boundaries is the lack of pervasive and high-quality bandwidth and, in the case of disadvantaged regions of the world, the lack of connectivity altogether. The proposed Postmanet has the potential of providing a form of inexpensive, pervasive, and plentiful bandwidth to the mass, potentially alleviating the "digital divide" phenomenon. (2) In terms of advancing our knowledge of the science underlying collaborative work, a substantial portion of this proposal is devoted to the understanding and development of collaborative peer-to-peer systems and applications built on top of the Postmanet. (3) An important Postmanet demonstration application area that we plan to target is basic education for disadvantaged areas. Due to its cost, bandwidth, and reach advantages, the Postmanet is well-suited for distance learning applications at places that need them the most: developing regions, for example, where illiteracy rates are high, the shortage of qualified teachers is severe, and resources are scarce. (4) More generally, we believe that the application possibilities enabled by an inexpensive communication channel that has a tremendous amount of bandwidth can be vast, and the Postmanet utility should be applicable beyond the disadvantaged regions.

Contents

C Project Description

C.1 Introduction

C.1.1 The Digital Divide

As the Internet continues to mature, latest studies [9] have found that there is significant increase in the usage of the net by people in their daily activities, in areas such as education and training, health care, politics, functioning of governments, religion, news, arts and entertainment, finance, and transaction-based activities. The Internet is literally continuing to transform every aspect of the society. The increasing adoption of broadband connectivity is particularly instrumental in opening up new application opportunities. Despite this growth in activity, however, the same studies also find that the growth of the online population has slowed down. There was almost no growth over the course of 2002, and the online U.S. adult population remains at sixty-three percent. Thirty-one percent of Internet users who go online from home have broadband as of August 2003. Despite the recent dramatic growth in broadband adoption, these studies also show that this trend is likely to slow down significantly as well [6].

As the adoption rates of the Internet and broadband connections slow down in the U.S., these studies find that the "digital divide" could be solidifying: those with modest incomes, rural residents, and minorities are among those who lag behind in Internet access. Most people living in developing regions, who represent an overwhelming majority of the world's population, also largely fall on the "wrong" side of the digital divide and are missing out on many important opportunities that the digital world could have provided.

Bridging this digital divide, especially by attempting to increase the accessibility of broadband connectivity, can be very challenging. The improvement of wide-area Internet bandwidth is constrained by factors such as how quickly we can dig ditches to bury fibers in the ground. The cost of furnishing "last-mile" wiring can be prohibitively high, and the progress has been excruciatingly slow.

C.1.2 Using the Postal System to Extend and Complement the Internet

In this proposal, we explore the use of digital storage media (such as DVDs or even hard disks) transported by the postal system as a general digital communication mechanism. While the idea of sending digital content via the postal system is not a new idea – companies (such as AOL.com and netflix.com) have used this approach to deliver software and movies on a large scale for some time – none of these existing attempts have turned storage devices delivered by the postal system into a *generic* communication channel that can cater to a wide array of applications. We shall call such a system a *Postmanet*. Compared to traditional wide-area connectivity options, the Postmanet has several important advantages.

• *Wider reach.* The postal system is a truly global "network" that reaches a far greater percentage of the world's human population. It is a robust, proven technology. It exists and works well today—no significant new investment in exotic equipment is required.

• *Greater bandwidth potential.* If we compare the number of bytes that can be shipped by the Postmanet against that can be transmitted over the traditional wide-area Internet in the time interval of one or a few days, it is a well known phenomenon today that the former can be far greater than the latter. Some may consider this phenomenon a temporary fluke as a result of the relatively poor capacity of today's Internet. We, however, believe that this is not the case. Our belief stems from an observation of some fundamental technology trends. Storage density has been increasing at the annual rate of 60% to 100% for many years, and it is likely to continue in the foreseeable future. As a result, the amount of information that can fit into a fixed amount of volume, or the amount of information that can be shipped by the postal system for a fixed cost, increases exponentially at roughly the same rate, a rate that the wide-area Internet bandwidth growth is unlikely to be able to keep up with. Indeed, far from being a temporary fluke, the bandwidth gap between these two modes of transport is only going to widen as the storage density continues its rapid improvement.

• *Low cost.* The goal of providing citizens with affordable access to postal service is typically an integral part of most nations' postal system charters. In the U.S., even if each house sends (and receives) a DVD each day, the monthly cost of about $10 compares favorably with existing ISP offerings, especially if we were to consider its vast bandwidth potential. The relative liberal use of the postal system by AOL.com and netflix.com highlights the low cost advantage

1

of this approach.

● *Better scalability.* The postal system is highly decentralized and does not appear to suffer from potential bottlenecks such as backbone links on the Internet. The postal system has tried and tested experience dealing with "flash crowds" such as those seen during certain holidays.

We note that our goal is *not* to compete against or to replace traditional Internet access; instead, our goal is to *extend* and to *complement* the Internet.

● *Extending the Internet.* For those who have no access to connectivity or high-bandwidth connectivity, the Postmanet can provide an inexpensive connectivity alternative to enable certain networked applications, especially bandwidth-intensive ones. The target audience may include not only rural residents, those who live in developing regions, and economically disadvantaged people, but also users such as business and recreational travelers to places lacking connectivity.

● *Complementing the Internet.* Although the Postmanet can yield enormous bandwidth, it has long (but reasonably predictable) latencies, such as a small number of days. We call such a channel a High Latency High Bandwidth (HLHB) channel. Correspondingly, we call a traditional Internet connection a Low Latency Low Bandwidth (LLLB) channel. For places that have access to both an HLHB channel and an LLLB channel, an interesting problem is to exploit an integrated and simultaneous use of *both* channels to get the best of both worlds. For example, small requests, acknowledgements, "NAKs," and control messages can be sent along the LLLB Internet, while large messages are staged on mobile storage devices for transmission by the HLHB Postmanet. Another example of the complementary nature of the Postmanet is that it may increase the availability of the communication subsystem: if the Internet is down for some reason, one still has another alternative.

At a first glance, it may appear deceptively simple for an end user to manually burn and ship DVDs for whatever applications he desires. But with many applications, hundreds of communicating parties, and thousands of "messages," for example, manual management would become infeasible. Therefore, systematic transparency support at the levels of network transport, programming models, routing, and applications is necessary. In Section C.2, we describe the user experience with the Postmanet and some of its example applications. In Section C.3, we describe systems support at the communication end points, which attempts to transparently emulate a traditional network on top of the postal system. In Section C.4, we examine options of "routing" data from the sender to the receiver in the Postmanet system: we shall see that if we were to scale the Postmanet to accommodate a large number of participants, we need better answers than "leaving it to the postal system." We discuss how a large-scale peer-to-peer system can be built on top of the Postmanet. In Section C.5, we discuss how we handle interaction in a key target application area of the Postmanet: the delivery of basic education in disadvantaged areas. We shall see two recurring themes at these different levels of the system. One is the simultaneous exploitation (discussed above) of the Internet and the Postmanet so we can combine their latency and bandwidth advantages. The other is the exploitation of the abundant capacity and bandwidth of the Postmanet to improve its latency, reliability, and cost. We describe related work, intellectual merit and impacts, and proposer's prior NSF work, in Sections C.6, C.7, and C.8 respectively.

C.2   Usage Scenarios and Applications

C.2.1   A Transparent Postmanet Channel

An important goal of ours is to make the Postmanet as transparent as a conventional network channel for a user. One way of better understanding this transparency is to visualize a box that is similar to a small conventional home network router: it allows several home computers to share a wide-area Internet connection. A Postmanet router is just such a box plus one or more DVD slots in the box (see Figure 1). During the day, a user of the Postmanet router simply uses his applications in a way that is entirely oblivious of the presence of the Postmanet. At the end of the day, the Postmanet router box automatically spits out an outgoing DVD filled with some data. A postman makes a routinely scheduled stop to pick up the DVD for delivery. (We will discuss options of postal label generation later.) Each day, the user also picks up an incoming DVD dropped off by a daily postman visit. The user does not have to manually inspect or process the content of the DVD in any way: he just inserts the DVD into a slot in the Postmanet router box. From that point on, the user continues to use his applications in an oblivious way.

The details of this imaginary Postmanet router box can vary. The box may not necessarily be a dedicated physical
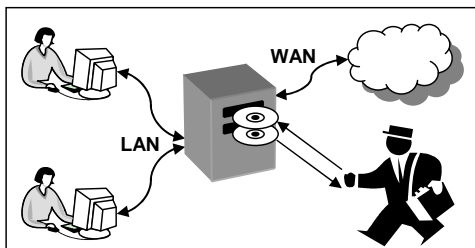
Figure 1: An imaginary Postmanet router box.

device: the user home computer may shoulder the task. We may use different types of mobile storage devices: these may include read-only or read-write DVDs, IBM Microdrive-like small hard disks, or flash memory cards. The number of the mobile storage devices picked up and dropped off by the postman per visit may vary. The Postmanet router box may or may not be complemented by a conventional wide-area network connection. The box may or may not be shared by multiple users. While all these details may vary, a constant is the transparency feature: the fact that the user's direct manual interaction with the box is limited to the insertion and removal of a couple of mobile storage devices per postman visit.

C.2.2 Example Applications

The following example applications share at least two common themes: (1) their bandwidth demands can far exceed those that can be met by a traditional wide-area network; and (2) these applications can benefit from the *simultaneous* exploitation of the HLHB Postmanet and the LLLB Internet.

• *Email with large attachments.* For example, one may be able to send large home movie files via email. This application may take advantage of the LLLB Internet by sending the small message body over it, while the large attachments travel over the HLHB Postmanet. (Certain extra UI features are needed to deal with the decoupled arrival.) This is one of the simplest applications of the Postmanet because the semantics of the email application are such that there is little distributed data coherence issue. While the traditional email application is primarily push-based, we could extend it to include a pull-equivalent of the attachment feature: subject to protection checks, a request for a large data item can be sent over the LLLB Internet, triggering an automatic email reply whose large data attachment is sent back over the HLHB Postmanet.

• *Web pointing to or embedded with large data objects.* These large data objects may include audio, video, programs, and models. To avoid the need of client-side browser modification, one can employ a Postmanet-aware client-side proxy. Small data items, such as top-level html pages, can be retrieved over the LLLB Internet (to ensure their freshness). Large data objects that are pulled or pushed over the HLHB Postmanet can be placed in a client-side cache. The client-side proxy may poll over the LLLB Internet to check the freshness of the cached data. There are two possibilities for the server side: the content publisher is either Postmanet-aware or not. A Postmanet-aware content publisher program may respond to client "subscription" requests by sending them large data items over the Postmanet. For a content publisher site that is not Postmanet-aware, a possible way for its large data items to reach poorly-connected clients is via a well-connected third party that *is* Postmanet-aware: this third party would retrieve large data items from the original content publisher over a conventional network and repackage them to send to poorly-connected subscribing clients over the Postmanet.

• *Remote file system mirroring for sharing and/or backup.* Large amounts of newly written file data can be transmitted to a remote mirror site over the HLHB Postmanet. Users at this remote mirror site who happen to desire to read up-to-date versions of the files may use the LLLB Internet to check freshness of the mirror.

• *Peer-to-peer file sharing.* The large sizes of media files make it appropriate to transmit the files over the HLHB Postmanet. The small messages generated by foreground file search processes or background announcements of sites' contents can be transmitted over the LLLB Internet. Note that the use of the Postmanet is orthogonal to the choice of the overall file sharing system architecture, which can be based on either centralized metadata servers or entirely decentralized alternatives. Copy right protection concerns can be addressed by incorporating Digital Rights Management (DRM) techniques [1] and the Postmanet should be neutral to such concerns, just as a traditional network

3

is.

● *Video "almost on-demand."* A shortcoming of the existing online DVD movie rental businesses is the multi-day latency elapsed between the time a request is submitted over the Internet and the time the desired movie arrives via the postal system. In an alternative model, subject to customer permission, the rental company could proactively push encrypted movies to participating customers without necessarily having received explicit requests. These may include recommended movies based on customers' rental history, popular movies, and new releases. At the current rate of storage density and price improvement, it would be very reasonable to assume the use of multi-terabyte hard disks filled with hundreds or even thousands of movies being employed by the HLHB Postmanet. Huge encrypted libraries of movies can accumulate on participating customers' local storage devices. To view a movie, a customer would purchase a decryption key on-demand from the rental company over the LLLB Internet and gain access to a locally stored and encrypted selection instantaneously. Again, emerging DRM technologies such as Microsoft's Palladium [1] should be able to prevent unauthorized dissemination of decrypted content or other usage that is outside a contract. Such a DRM contract may restrict, for example, the number of times that a movie can be played for a certain payment.

● *Publish/subscribe systems for other types of content.* The above video "almost on-demand" application can be generalized to disseminate many other types of content in a generic publish/subscribe system. The types of content may include music, TV and radio programs, newspapers and magazines (with richer presentation), store catalogs (with richer presentation), software releases and updates, and public lectures given at universities. The possibilities enabled by an inexpensive communication channel that has practically infinite bandwidth can be vast. While it is possible to develop and maintain individual solutions for different types of content, the presence of a *generic* Postmanet infrastructure that is available to all applications makes the approach more attractive and easier to manage for all types of content publishers and consumers.

● *Distance learning.* In addition to multimedia teaching material that is being disseminated by teacher sites to students over the HLHB Postmanet, the students may submit content such as digitized homework for grading over the Postmanet, and the resulting teacher feedback may be sent back over the Postmanet again. Smaller network messages such as teachers' synchronous commands controlling the real-time playing of teaching material at students' sites may be transmitted over the LLLB Internet. We will explore this application in greater details in later sections.

In the above discussion of example applications, we have explained how we can exploit the simultaneous use of the HLHB Postmanet and the LLLB Internet. The availability of an LLLB Internet, however, is not absolutely necessary for the functioning of the HLHB Postmanet, or for demonstrating its power. For example, a user may receive a large digital catalog of Amazon.com via the Postmanet, browse the catalog and place purchase orders "off-line," and send back the orders via the Postmanet (instead of via an Internet connection, had it been available). This arrangement is especially useful for places such as isolated remote regions in developing countries where even dialup connections are not always available. (Yes, even the economically disadvantaged have purchasing power, and it is often useful being able to transact with sources outside of their confined neighborhoods that may offer limited choices and/or artificially inflated prices[10].)

Although we have called the Postmanet a "high latency" channel, we note that by exploiting the plentiful storage capacity and bandwidth of the Postmanet, it can be possible to mask its high latency. The video "almost on-demand" example is particularly illustrative: by liberally disseminating content that may never be actually used by users who receive it, a publisher who uses the HLHB channel can, perhaps ironically, create the illusion of instantaneous on-demand access for content that *is* used. This theme of deliberately "wasting" plentiful resources to optimize for scarce resources will be revisited in other aspects of the functioning of the Postmanet.

C.3 End Point Support for Transparency

C.3.1 Complexities of Manual Management

At a first glance, the idea of burning and shipping DVDs seems straightforward enough for end users to manage manually. In reality, however, it can be more complicated than it may appear, if we were to make the most out of this approach. In this subsection (Section C.3.1), the terms "sender" and "receiver" refer to the two *human* users at the two ends of a Postmanet communication channel.

A sender would have to consider which one of the Internet and postal system channels to use for a particular data

item. If the postal system is chosen, the sender would have to assign some meaningful names to the data items that are to be placed on outgoing mobile storage devices. If these data items are rewritten before they are shipped, the sender would need to ensure that the proper versions get placed on the mobile storage devices. The sender would have to prepare instructions for the person who is to receive the mobile storage devices via the postal system. The sender would have to manually copy data onto the outgoing mobile storage media. The sender would have to worry about whether the storage media could have been delayed, damaged, or lost in the postal system. To each of these possibilities, the sender would have to know how to systematically respond. The sender may need to worry about ensuring the privacy and security of the data contained in mobile storage devices. The sender may need to query the intended recipient about shipping status. The sender may need to respond to queries from the recipient about shipping status and instructions for handling the arriving storage media.

The tasks facing a receiver of the mobile storage devices delivered by the postal system can be equally unpleasant. The receiver would have to decipher the instructions on what to do with the arriving data. The receiver would have to arrange for acknowledgements to be sent back. The receiver would have to manually copy data out of the arriving mobile storage media. The receiver would have to worry about out-of-order delivery of multiple mobile storage devices by the postal system. The receiver would have to worry about out-of-order delivery of data arriving on the Internet with respect to data delivered by the postal system. For data that spans multiple mobile storage devices, or data that is partially sent over the Internet and partially sent via the postal system, the receiver would need to know how to put them back together. The receiver would have to worry about answering queries from the sender about shipping status. The receiver may need to check for possible tampering of data. If the mobile storage media arrives damaged, the receiver needs to know how to respond.

And perhaps worst of all, for both the sending and receiving sides, the lack of systems support may imply that there is not a programming interface for applications to programmatically and easily exploit, let alone to coordinate or share the use of this communication mechanism among them. Clearly, without automatic systems support, the utility of this approach would be severely limited.

C.3.2 Sender

One of our goals is to automate away almost all of the manual tasks described above. The human user does not manage any low-level details such as naming temporary data files on the mobile storage media. A human user may choose to provide some hints or specifications in terms of, for example, at what time in the future he desires the data to arrive at the intended recipient. In this subsection and the next (Sections C.3.2 and C.3.3), the terms "sender" and "receiver" refer to the system softwares residing at the two ends of a Postmanet communication channel.

The sender does not necessarily have to worry about which of the Internet and Postmanet channels to use to meet the users' needs; the system could make a choice based on, for example, the amount of data that needs to be sent and the desired arrival time. Indeed, the system may choose to use the Internet and Postmanet channels in parallel. Portions of a large data object may start to incrementally arrive at the receiver over the Internet, while the complete object arrives later over the Postmanet. At the application level, multiple versions of a data item may be prepared: a low resolution version is shipped over the LLLB Internet, while a high resolution version is shipped simultaneously via the HLHB Postmanet. Multiple versions of the data may "race" against each other as they progress in the two different "networks," so we can trade off metrics such as quality, latency, and availability. If the Internet is "max'ed out," the system needs to carefully prioritize what is sent through the Internet connection. One way of looking at this problem is to view the Internet connection as a "cache" of the Postmanet connection: the former is a faster (latency-wise), smaller (bandwidth- and capacity-wise), and sometimes more expensive alternative that provides comparable functionalities.

In addition to possibly sending redundant data simultaneously over the Internet and the Postmanet, we may also proactively replicate data in the Postmanet. For example, as multiple mobile storage devices are sent between a sender-receiver pair on successive days, we may liberally replicate outgoing data of earlier days on outgoing devices sent on later days. In cases where a single storage media is delayed or lost due to accidents in the postal system, the replicated data on subsequently arriving devices is just a day away, so we can avoid unnecessary long end-to-end retransmission delays. This is another example of the consistent Postmanet theme of liberally "wasting" plentiful resources (storage capacity) to optimize for more difficult metrics (low latency).

The system would automatically perform the necessary copying to prepare mobile storage devices for shipment. At the application level, if a later sending event should supersede earlier ones (because, for example, only the freshest

version of an updated object needs to be sent), the system would take care of excluding from the mobile storage device data that is sent earlier. The system would multiplex data sent by multiple applications and/or multiple users onto one or more outgoing storage media. The number of storage devices that are needed depends on not only the total size of data, but also the scheme that we use to forward data in the Postmanet (as we explain later). If the mobile storage devices being used can be written to incrementally, we may not need to wait till shortly before the arrival of the postman to begin writing to the device in a long burst—continuous background copying could have occurred throughout the day. At the application level, as we have noted earlier, due to the relatively large capacity of storage devices, as long as there is enough time to perform the writing/copying, the system should be very liberal in deciding what gets placed on the outgoing storage devices. For example, even when there is only a very small chance that a piece of data will eventually be used by the remote recipient, it does not hurt to go ahead and copy the data anyhow. The sender may also compute fingerprints and/or encrypt data to ensure tamper-resistance and/or privacy.

The sender should not have to worry about ad hoc monitoring or querying of transmission progress. In addition to receiver acknowledgements, the postal system may provide online tracking (transmitted over the Internet) that aids the sender's progress monitoring. If the sender fails to receive a receiver acknowledgement (which should also arrive over the Internet) after a certain period of time (due to possible loss of the mobile storage media by the postal system), or if the sender is informed by the receiver that the data is corrupted (due to a damaged storage device), it may take one of several possible actions. If there is a previously proactively transmitted replica of the data already in transit, the sender informs the receiver and does nothing else. If not, and there is enough time left before the data is to be used by the receiver, the sender may "retransmit" by placing a new copy of the data onto a new mobile storage media that is to be picked up by the next scheduled postman visit. If there is not enough time for retransmission in the Postmanet, the sender system software may invoke some type of application-specific cooperation. For example, the application may provide a much lower-resolution version of the lost or damaged data for retransmission over the Internet.

C.3.3 Receiver

The arrival of the data at the receiver can trigger several types of responses. The receiver first needs to send back an acknowledgement over the Internet. The receiver may copy data from the incoming mobile storage media into a local store. (The Postmanet router box shown in Figure 1 may include a hard drive.) The receiver may also trigger application-level activities. For example, for the email application, the arrival of a large attachment may trigger UI responses in the recipient's email reader.

There can be a number of systems-level and application-specific reasons that make it unnecessary to copy the data from an arriving mobile storage device. At the systems level, for example, due to proactive replication or premature retransmission by the sender, either across multiple mobile storage devices, or across the Internet and the Postmanet, the receiver may need to discard the duplicates. At the application level, an application can be granted direct access to the data contained on the mobile storage device to make data copying out of the mobile storage device unnecessary. For read-only storage media, copy-on-write techniques may be necessary. Multiple storage media may have been delivered by the postal system out-of-order, so some of the data on a late-arriving device may be stale and is superseded by data on an earlier device. Similarly, the freshest data that is sent later may have arrived earlier over the Internet, which again renders some of the data on the earlier sent mobile storage device obsolete. Even in absence of fresher data at the receiver, the receiver application may discover that some of the newly arriving data is no longer needed due to other reasons. In all these cases, the system must exercise care not to unnecessarily copy or use obsolete data. This is especially relevant for the Postmanet due to a general Postmanet theme of liberally "wasting" storage capacity by the sender as long as there is time to do it in an effort to optimize for other metrics—this "wastage" by the sender, however, needs to be checked with receiver intelligence.

The receiver needs to play its part in the security mechanisms. When data arrives in the Postmanet, the receiver needs to ensure that (1) the data is from a sender whom it is willing to receive data from; (2) the sender identity is not forged; and (3) the data has not been tampered with. If desired, a version of this system may be more liberal in its policy of accepting data so, for example, some type of "spam" could be allowed but given a lower priority of receiver processing. (In some scenarios, indiscriminate spam may be less likely to be a serious problem when the sending side must pay the postal system to get his message across.)

C.3.4 Observations

Many of the issues described above, such as retransmission, handling out-of-order delivery, suppressing duplicates, and minimizing data copies, bear a resemblance to those that one must deal with in traditional communication networks. In the context of the Postmanet, however, the boundary between storage and networks is blurred. The latency and the amount of data involved in a "packet" (i.e., a mobile storage media) can be many orders of magnitude greater than those of a traditional network packet. This makes the research problem as much a distributed storage problem as a networking problem.

In some application scenarios where data coherence is relevant, for example, the freshest data can be distributed over a number of devices: the local stores at the receiver and sender, the mobile storage devices being prepared by the sender, the mobile storage devices that have arrived at the receiver, and the devices that are in-transit. The application may need to "know" where all the pieces are so it can put all the jigsaw puzzle pieces together without physically copying all the pieces to one place.

Two recurring themes are also unique to the Postmanet work, both at the application level and at the systems level. One is liberal use of available storage capacity. For example, as we have described, at application level, we may choose to freely transmit many data items over the Postmanet that have only small chances of getting used; and at the systems level, we may choose to replicate freely the same data items. At both levels, the benefit can outweigh the cost. The second theme is the judicious simultaneous exploitation of the LLLB Internet and the HLHB Postmanet. For example, as we have described, at the application level, user-visible objects of different sizes are placed on one or both of the two "networks;" and at the system level, we apply the same principle to control messages that are not visible to users.

C.3.5 Programming Models

We have focused on the simple Postmanet data transport primitives above. It may be desirable and necessary to synthesize higher-level programming models to make the use of the Postmanet by programmers easy. For example, consider the implementation of asynchronous "reads." A read request is sent over the Internet. Upon arrival, if the desired data item is large, the read request would automatically trigger the desired data to be staged on a mobile storage device, which would be transported back over the Postmanet. When the desired data arrives at the requester, it may trigger yet another action that the requester has pre-specified. This style of communication is similar to some existing asynchronous communication models [16] and the programming languages built on top of them [2]. Under this programming model, handler codes attached to messages are asynchronously executed upon arrival of the messages, to incorporate the newly arriving data into ongoing computations. Investigating the applicability and developing extensions of such asynchronous communication programming models is a key element of our proposed research. In our description of the end point mechanisms in Sections C.3.2 – C.3.4, we have used phrases such as "at the application level" and "at the systems level;" many of these application-level strategies could use support provided by an application programming interface (API).

C.4 Routing

So far, we have considered system support at the communication end points. We now consider how a mobile storage device is routed from the sender to the receiver. One may naively think that the postal system would take care of it all, but as we shall see, the issues in the Postmanet can be more complicated, especially if there are a large number of communicating parties.

C.4.1 An Analogy

We start by considering the netflix.com movie DVD rental business as an analogy. In the early days of the company, when customers return their DVDs, they are always first shipped back to the company headquarters in San Jose. The disadvantage of this approach is obvious. For example, when an east coast customer A returns a DVD that A's next-door neighbor B is waiting for, despite the close proximity of A and B, the DVD is forced to make a long detour through the west coast.

Nowadays, netflix maintains multiple distribution centers throughout the country. A returning DVD is sent to the distribution center that is closest to the customer's home address. Upon the arrival of the DVD, the staff at this
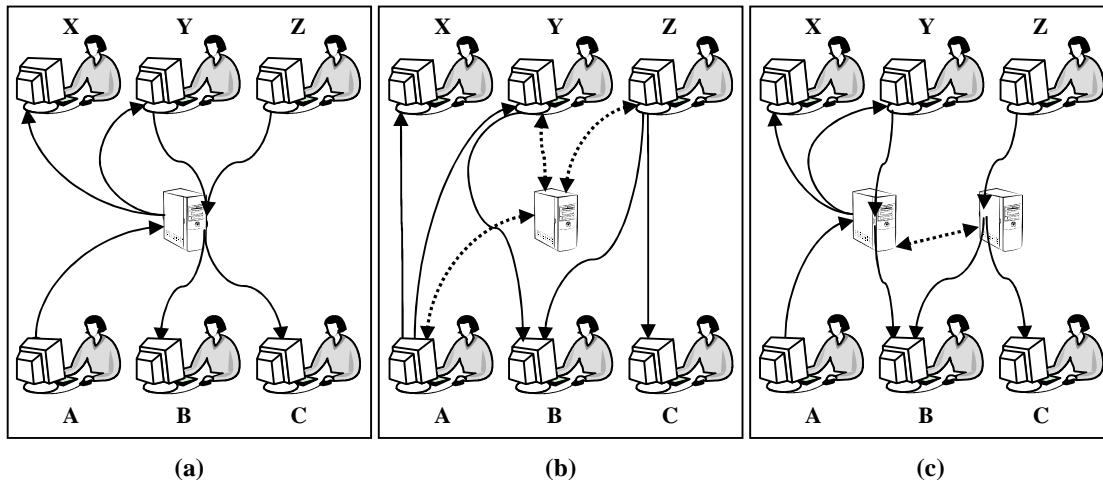
Figure 2: Three simple routing strategies. A solid arrow denotes data communication carried by the Postmanet. A dotted arrow denotes routing information carried by the Internet. In all three panes, A sends different data items to X and Y, Y sends some other data to B, and Z sends to both B and C. (a) Centralized data routing via a single data distribution center. (b) Direct peer-to-peer data routing. (c) Data routing via multiple data distribution centers.

distribution center looks up a database of waiting lists to see if any other customer is waiting for this DVD. If there is, the DVD is then forwarded immediately to the waiting customer at the head of the queue. This improvement has significantly reduced the wait time for netflix customers and unnecessary extra costs associated with long hops in the postal system. This analogy is illustrative, but the issues in the Postmanet are even more complex: the data granularity and communication pattern in a movie rental business are much simpler than those seen in a generic communication mechanism.

C.4.2 Three Simple Routing Strategies

Let us consider the examples illustrated in Figure 2. Figure (a) is the centralized alternative. While the disadvantage of this approach is obvious, it has an important advantage. For example, even though A needs to send data to two receivers, A only needs to send a single mobile storage device to the central server, which acts as a "switch." Similarly, even though B needs to receive data from two senders, B only needs to receive a single mobile storage device from the central server. The mailing labels used by all the end communicating parties are identical. The server copies data from incoming storage devices to outgoing devices, so that each site at most receives one storage device and sends one for each postman visit. In effect, data routing occurs both digitally and mechanically: digitally when data is copied from one storage device to another at the central server, and mechanically when a storage device is carried to and from the server by the postal system. The server also enforces security policies so, for example, only data that originates from authorized senders and is untampered with is forwarded.

Figure 2(b) illustrates an "opposite" approach. The role of the central server is limited to the coordination of routing decisions: it does not participate in data forwarding. In this figure, A consults the central server to obtain information such as the mailing labels of X and Y. A then sends two separate mobile storage devices directly to the intended recipients. (It is interesting to speculate how a peer-to-peer version of netflix.com may operate based on this approach.) Data routing is potentially more efficient than that in Figure (a). On the other hand, a disadvantage of this approach is that a site could receive or send a large number of storage devices per postman visit, which could become an administrative and cost burden. For a modest-sized system, however, this approach can be an attractive approach as it demands the least from a shared infrastructure.

In Figure 2(c), we employ multiple netflix-like distribution centers that are geographically distributed. As is the case in Figure (a), each site needs to send at most one storage device toward the closest data distribution center per postman visit. Unlike the simpler netflix case, data may need to be recopied among devices at the distribution centers. Each site may receive multiple devices per postman visit, as many as the number of distribution centers. (Or alternatively, a site may send multiple outgoing devices but receive only one incoming device per postman visit, if
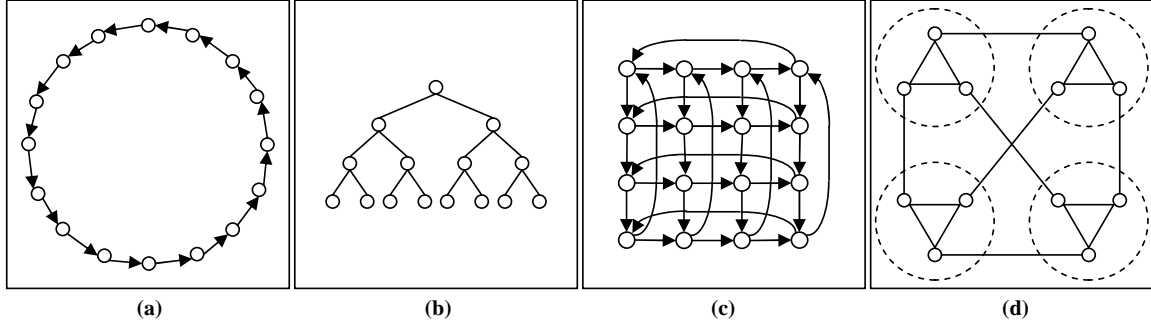
8

Figure 3: Indirect peer-to-peer routing topology graphs. Each node represents a Postmanet site. A directed edge indicates that data can only flow in the direction of the arrow. An undirected edge indicates that data can flow in either direction.

we insist that a mobile storage device must be sent to the distribution center closest to the receiver, not the sender. Or alternatively, a site may employ a mixture of these approaches and send and receive multiple devices per postman visit. In all cases though, the number of devices involved per postman visit is limited by the number of distribution centers.) The geographically distributed distribution centers allow some degree of geographical awareness in routing decisions. The distribution centers do not exchange data with each other, but they may communicate among themselves to coordinate routing decisions. The latency achieved under this alternative is likely to be worse than that is possible under the alternative illustrated in Figure (b) due to the extra hops through the distribution centers. It is possible to allow the coexistence of the alternatives illustrated in Figures (b) and (c), so occasional latency-sensitive packages can be routed directly to their destination without passing through data distribution centers.

C.4.3 Static Indirect Peer-to-Peer Routing Strategies

The routing strategies that we have examined above have disadvantages. The approaches illustrated in Figures 2 (a) and (c) utilize data distribution centers, which can be a non-trivial infrastructure investment if these servers need to copy a large amount of data for many Postmanet end users. Under the direct peer-to-peer routing strategy shown in Figure 2 (b), if there are many Postmanet users, each user may need to send and receive many mobile storage devices per day. A teacher who returns graded homework in a distance learning system, for example, may need to send as many DVDs as the students that she has.

An ideal Postmanet routing mechanism should possess the following characteristics: (1) it can accommodate a large number of simultaneous Postmanet communicators without requiring a user to handle many mobile storage devices per postman visit; (2) it has end-to-end message propagation latencies that are close to those provided by the postal system; (3) it does not require an expensive infrastructure other than the existing postal system; and (4) it does not burden Postmanet nodes in an unbalanced manner with data copying tasks that are beyond their own communication needs.

By requiring Postmanet user nodes to forward data destined for others, we may be able to, in some sense, distribute the data copying tasks of a data distribution center infrastructure among the end user participants so we can eliminate the need for such an infrastructure, thus combining the advantages of the different strategies shown in Figure 2.

Figure 3 shows some ways that Postmanet nodes can forward data for each other. Suppose the number of Postmanet nodes is N. In the following discussion, when we say a user "handles" k disks, we mean that the user may receive up to k disks and send up to k disks per postman visit; and when we refer to a "latency" metric, it is in terms of the number of postal system forwarding hops visible to Postmanet participants. Using the ring topology of Figure (a), a Postmanet user handles one disk but the end-to-end latency can be long. Using the tree of Figure (b), a Postmanet user handles three disks and the end-to-end latency is $\log N$, but the higher-level nodes are likely to need to copy more data. The ring topology of Figure (a) can be extended to a 2-d mesh as shown in Figure (c), or more generally, to a k-dimensional mesh. In such a topology, a Postmanet user handles k disks and the worst case end-to-end latency is $k\sqrt[k]{N}$. Figure (d) illustrates a hierarchical cluster-based algorithm for constructing a routing topology: an "$(i + 1)$th level cluster" is a complete graph of $j + 1$ lower-level (ith level) clusters, each of which has j nodes inside. In Figure (d), for example, $i = 1$ and $j = 3$. At the next two levels up, j grows to $3 \times 4 = 12$ and $12 \times 13 = 156$, respectively. In this topology,

9

the worst case end-to-end latency is $O(\log N)$, and the number of disks handled by each user is $O(\log \log N)$.

Other alternatives include the various interconnection topologies used by Distributed Hash Tables (DHTs) to enable object location in peer-to-peer systems. These include adaptations of hybercubes [12, 19], tori [11], and de Bruijn graphs [8]. The hybercube-based interconnections, such as Pastry and Tapestry, can be utilized in a Postmanet to probabilistically limit the end-to-end latency to $O(\log N)$ with each user handling $O(\log N)$ disks. Topologies based on de Bruijn graphs can also probabilistically limit the end-to-end latency to $O(\log N)$ while requiring only a constant number of disks handled by each user. Many of these systems also exhibit fault tolerance as they are able to successfully route even if a certain fraction of the nodes fail. Furthermore, these systems employ implicit routing wherein routing decisions are made locally without requiring elaborate knowledge of the global topology. We do however note that implicit routing is of limited value in Postmanet, where the control and data traffic can be conveyed on different networks – the LLLB Internet could be used for dispersing topology information while the actual data is communicated over the HLHB channels.

In graph theoretic terms, this problem of designing a good interconnection network can be seen as that of constructing a directed graph with a large number of nodes while keeping the diameter and the maximum node degree small. The diameter corresponds to the maximum end-to-end latency as messages are forwarded along the edges of the graph, and the degree of a node corresponds to the number of disks it handles. For diameter D and maximum degree Δ, the following can easily be shown to be an upper bound (known as the *Moore bound*) on the number of nodes in the graph.

$$1 + \Delta + \Delta^2 + \ldots + \Delta^D = \frac{\Delta^{D+1} - 1}{\Delta - 1}$$

It can be shown that the Moore bound can not be achieved exactly, except for the trivial cases when Δ or D is 1. There is, however, a long history of attempts to explicitly construct large graphs with bounded diameter and maximum degree. (See [4] for example.) Many of these constructions use sophisticated group theoretic techniques. Randomized constructions with constant degree for each node and $O(log\ N)$ diameter are also known. We, however, note that it is important to map these abstract graphs onto Postmanet participants in a geography-aware manner so that we do not incur unnecessary extra postal system delays and costs. Therefore, simpler and more intuitive constructions, such as those illustrated in Figure 3, may eventually prove to be more usable in practice.

C.4.4 Dynamic Peer-to-Peer Routing

The possible routes described in Section C.4.3 are static: a Postmanet node may communicate directly only with a small number of pre-determined "neighbors." These static constraints may be unnecessarily restrictive. For example, if A sends to B, and if neither A nor B is communicating with or forwarding data for anyone else, there is no reason why A should not be allowed to send *directly* to B. The goal of a more dynamic approach is to allow for such routing flexibilities without causing problems such as too many disks being handled by any one node on any given day. This dynamic routing problem is made easier by the fact that we may potentially use the LLLB Internet to exchange traffic and routing information among the Postmanet nodes. At the same time, this problem is made more difficult by the long latencies of the postal system: for example, when a disk is sent on Monday from A to B and is expected to arrive on Thursday, it may be difficult to accurately predict on Monday how many disks B would receive on Thursday.

Another strategy that we may consider is to replicate data and send them down multiple postal routes simultaneously. Such copies can be had for free (in terms of postage and other costs) if these multiple routes would have been used anyhow even in absence of the replication strategy. Not only may this strategy improve latency and reliability, it can also increase routing flexibility as some replicas can be freely discarded in the middle of some routes. We may use the LLLB Internet to "shoot down" extra replicas, for example, when one of the copies reaches the destination, or more generally, when it becomes obvious that the further forwarding of these extra replicas would result in sub-optimal behavior. We have seen the themes of exploiting the LLLB Internet and exploiting the abundant storage capacity/bandwidth in the application (C.2) and end point management (C.3) sections; we now see the same themes recurring in the routing section.

While not technically falling under the title of "*dynamic* routing," another possible use of the Postmanet is to compose an end-to-end communication path with a sequential concatenation of Internet and Postmanet hops. For example, an isolated village may communicate with the rest of the world using a Postmanet hop, while the hops inside the village and the hops in the rest of the world are conventional network hops.
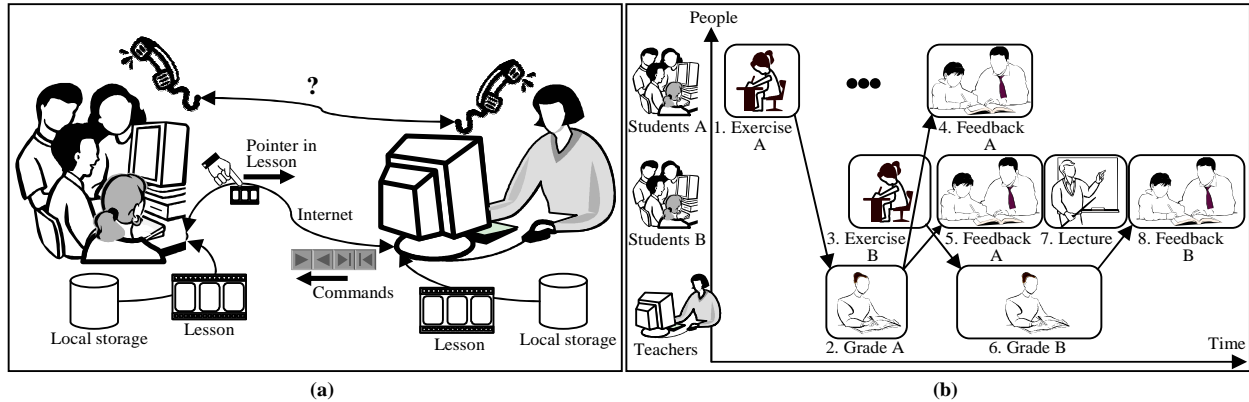
Figure 4: (a) Synchronous interaction. (b) Asynchronous communication: an example Gantt chart. The arrows denote communication, and the numbers denote the order of events.

C.5 Distance Learning Applications

An important Postmanet demonstration application area that we plan to target is basic education for disadvantaged areas [18]. It is important to note that our goal is not to compete against or replace human teachers; on the contrary, our goal is to amplify the reach and power of the limited number of qualified teachers that we do have. Today, remote disadvantaged areas, especially those in developing regions such as India (which has an illiteracy rate of 40%), tend to experience critical shortage of motivated and highly trained teachers. Distance teaching, which may allow flexible location and time commitments, can offer an attractive alternative to those who are enthusiastic about a career dedicated to helping needy children. And from the point of view of a child, although distance learning may not match the quality of a good face-to-face session, in many cases, it can represent a significant improvement over a dismal baseline. The Engineering School at Princeton University is to spearhead a new thrust aimed at researching and developing technologies that can improve the lives of people in less developed regions of the world. We expect our system to play an important part in this exciting initiative.

There are at least two main challenges facing a traditional distance learning system: (1) the need of providing sophisticated modes of interaction; and (2) the need of adapting to bandwidth constraints. To a large extent, these are conflicting goals, and existing approaches tend to favor one at the expense of the other. At one extreme, content is disseminated via TV broadcasts or on storage media (such as CDs or DVDs), and the amount of interaction is either non-existent or very limited. While such an approach may be useful for mature learners who are self-motivated and reasonably technically savvy, it may not work as well for young learners, who may need closer supervision and more personal and immediate interaction with teachers. At the other extreme, an existing approach to distance learning is similar to teleconferencing, which requires a large amount of communication bandwidth.

While the Postmanet may solve the bandwidth problem, what is less apparent is whether it is possible to provide meaningful interaction in a system built on top of it. It is this issue, *interaction*, that we focus on in this section.

C.5.1 Synchronous Interaction

Some of the interactions may have relatively low bandwidth demands, and they may occur in real time. Voice carried by phone or the Internet is an example. For a teacher to interact with the pupils effectively, she should know the context in terms of what the pupils are currently learning. For example, as illustrated in Figure 4(a), if a pupil asks a specific question about a subject matter in the middle of the content stream that has been transmitted to a local disk over the HLHB Postmanet ahead of time, the teacher needs to "see" that part of the content. In a traditional distance learning system, this interaction is accomplished through two-way video conferencing. In our system, since both the teacher and the students should have access to the same data stored on their respective local disks via the distance learning software implemented on top of the Postmanet, all the teacher needs are "pointers" into the content stream; these pointers are transmitted from the pupils along with their questions in real time over the LLLB Internet. In addition to providing voice feedback, the teacher can also send back computer commands over the Internet to control

11

the "playback" of the lesson at the students' site. These commands may rearrange the play sequence of a lesson, or even play a pre-canned response to an anticipated question.

C.5.2 Asynchronous Interaction

With limited bandwidth of the LLLB link, or the absence of one altogether, it is not always possible to provide sophisticated real-time interaction. Indeed, even in a conventional face-to-face classroom, from the point of view of any one particular child, the amount of real-time face-to-face interaction with the teacher is limited by the "bandwidth" of the teacher: the time of the teacher must be multiplexed across multiple pupils. A child, however, benefits from observing the interaction between the teacher and other kids. This indirect interaction, along with direct interaction, is what we seek to exploit.

Let us consider the example Gantt chart in Figure 4(b). In this figure, there are two groups of students: group A and group B. These two groups may or may not be co-located. The two groups perform the same exercises at time steps 1 and 3, respectively. The exercises are graded at time steps 2 and 6. And the resulting feedbacks are played in front of the students at time steps 4, 5, and 8. The time scale illustrated in the figure can range from a few minutes to several days, depending on the amount of data that is involved and the communication channel (the Internet or the Postmanet) that is used.

For a child in group B, he should see an uninterrupted stream of carefully prepared material. The stream includes instantaneous feedback (although the feedback consists of the teacher's interactions with other kids in group A) and specific feedback tailored for him (although the feedback is delayed). Consequently, the stream of events observed by a kid in group B should not be very different from what he would see if he were in a conventional classroom.

C.5.3 A Peer-to-Peer Learning Site

In subsections C.5.1 and C.5.2, we have mostly focused on point-to-point interactions. We now discuss many-to-many interactions. Our goal is to build a distributed peer-to-peer Postmanet-based learning site/system that brings together participants who are widely distributed in both space and time. As long as there is a quality control mechanism in place, we would like to allow participants to contribute in a flexible way that suits their time commitments and skill sets. In effect, we would like the learning system to act like an eBay-like "marketplace" where the demand and supply of various services are being matched up. (Note that whether people who provide services are offering them for free is an orthogonal issue: we anticipate both volunteer helpers who are offering their time for free and professionals who perform distance teaching for a living. The kids participating in the system should get their education for free.)

Participants may "plug" themselves into the system to play various roles. Some may develop new teaching materials and make it available for potential learners. Some may start new local schools to tap into an online (Postmanet) feed. Some students may choose to join as individual learners. Some may choose to serve as teachers. Some may choose to act as homework graders. Some may hold virtual "office hours" to provide students with additional help. In addition to student-educator interactions, the system also facilitates student-student and educator-educator interactions.

One of the relevant interesting technical aspects of such a site is that this is an application of the various Postmanet routing alternatives discussed in Section C.4: these alternatives show how a large-scale Postmanet system can be built with infrastructural support (in the form of one or more data distribution centers) or without.

More generally, this peer-to-peer learning system is interesting, compared to state-of-art peer-to-peer systems in the following aspects. First, the system exploits the use of the Postmanet. The recurring themes of exploiting the simultaneous use of the Internet and the Postmanet, and of liberally exploiting abundant storage/bandwidth capacity of the Postmanet, manifest themselves again at the level of exploiting the application-specific learning system semantics. For example, if a chosen remote grader of some homework assignments falls ill, it may take days for the postal system to deliver another copy of the assignments to an alternate grader. One way of addressing this issue is to liberally replicate data so that the homework submission is placed on the mobile storage devices sent to many teachers. (We have shown in Section C.4 that it is possible to do so without necessarily incurring even extra postage cost.) The first grader who chooses to grade the homework would send notification messages via the LLLB Internet to all others who have received (or will receive) replicas, so that no one else duplicates the effort. (Naturally, all the notification and cancellation mechanisms should be automated.) The second interesting aspect of this system compared to existing peer-to-peer efforts is that while there are many attempts examining lower-level file sharing and routing algorithms,

there have been few efforts on higher-level peer-to-peer collaborative applications.

Such a system can also encourage specialization, a principle in modern economies that promotes increased efficiency. Traditionally, teachers tend to shoulder many distinct duties. While the concentration of these distinct duties in single individuals has important advantages, it could also cause inefficient utilization of precious human expertise, a crucial resource that we try to optimize for. For example, the grading of simple forms of homework could be handled by individuals with relatively lesser skills or less experience. The best teachers should not be burdened with mundane tasks if we are to make the most out of their time. A distance learning system that brings many people together may be able to more efficiently schedule the human resources based on their skill sets and achieve a high degree of specialization.

C.6 Related Work

Due to space constraint, this section is not meant to be an exhaustive survey of all related work: we only selectively discuss some of the most relevant work.

C.6.1 Delay-Tolerant Networks

While the idea of sending data by using the postal system to deliver mobile storage devices is not new, and companies such as AOL.com and netflix.com have taken advantage of this approach to ship software and movies for some time, none of these existing efforts has turned the postal system into a *generic* digital data communication channel.

Recent efforts on "Delay-Tolerant Networks" (DTNs) [3, 5, 7, 13] have started to examine the use of WiFi-enabled mobile elements (such as buses equipped with storage devices) to simulate "delayed" connectivity to places that have access to none today. There are several important differences between existing DTNs and our proposed effort. While "postal classes of service" have been mentioned [3], to the best of our knowledge, the postal system has so far only been mentioned as an *analogy*—no existing system that we are aware of literally and explicitly proposes to exploit the postal system as a way of extending and complementing the Internet. Compared to existing DTNs, the literal use of the postal system presents several important different characteristics, most of which are positive; and these differences result in different research issues.

• *Better reach.* While existing DTNs are largely confined to relatively small regions or specialized environments, the postal system is a truly *global* "network" that reaches a far greater percentage of the world's *human* population. While much DTN research effort has been devoted to ad hoc routing, the postal system has its own proven and mature "routing" mechanism. While existing DTN proposals require investing in mobile equipment (such as WiFi- and storage-enabled buses), shipping storage media via the postal system already works fairly well today without investing in exotic equipment. While it is necessary to figure out how to bridge the multiple existing DTNs with the traditional Internet to provide end-to-end connectivity, and we still have not seemed to have a working solution yet, the postal system already provides end-to-end delivery today. Although it is possible to mix the use of the Internet and the postal system in various ways (as described throughout this document), and these strategies can be important optimizations, they are not necessities for obtaining end-to-end delivery. (See the imaginary Amazon.com scenario in Section C.2.2.)

• *Richer resources.* The existing DTNs are also frequently referred to as "challenged networks," which imply severe limitations on various resources, including the limited available communication bandwidth among mobile ad hoc elements, limited storage space on these nodes, and power consumption constraints. These limitations do not exactly apply to mobile storage devices shipped by the postal system. Indeed, as we have discussed throughout this proposal, one of the primary technology motivations behind this approach is the observation regarding its huge bandwidth potential. Mobile ad hoc DTN elements typically rely on WiFi-based wireless communication during periods of brief and/or intermittent contacts among these elements. In contrast, the mobile storage devices in the Postmanet are "dumb" and "dormant" during transit in the postal system. When they reach their destinations, they are "plugged in," quite possibly with high-bandwidth wired alternatives (such as USB2 or Firewire). Once such "contacts" are established, they may remain connected for extended periods of time. The typical DTN problems in bandwidth, storage space, or power consumption are not necessarily major concerns in the Postmanet usage model.

Due to these basic environmental differences, the research issues that we face are different from those facing existing DTNs. Issues such as ad hoc routing, flow control, congestion control, and managing buffer space and power consumption in mobile nodes are *not* necessarily our top focus. On the other hand, some of the research issues

discussed in previous sections are unique to our environment. For example, the relative abundance of storage capacity and bandwidth potential of the Postmanet leads to different optimization goals: instead of the careful conservation of these resources in typical DTNs, we may strive to "waste" some of the abundant resources in order to gain other advantages. Another unique aspect of our proposed system is the *parallel* use of the High Latency High Bandwidth channel (the postal system) and the Low Latency Low Bandwidth channel (the Internet). For example, small requests, acknowledgements, "NAKs," and control messages can be sent along the Internet while large messages are staged on mobile storage devices. Content of different degrees of resolution can be placed on both "networks" so multiple versions can race against each other as we trade off metrics such as quality, availability, and latency. In general, such techniques involved in the *parallel* exploitation of multiple connectivity technologies would be different from those involved in the *sequential* forwarding of data from one connectivity technology to another.

The larger research issues that we have described, including transport mechanisms, programming models, peer-to-peer routing schemes and applications suitable for the Postmanet, and applications centered around supporting distance learning, are also different from and/or complementary to the existing DTN research agendas.

C.6.2 Mobile Storage Systems

Despite the surface-level commonality of using mobile storage devices, the proposed Postmanet system is actually very different from existing mobile storage systems (some of which have been developed by us [14, 15, 17]). The goal of these distributed mobile storage systems is to provide the illusion of a coherent *disk* or *file* system, while the goal of the Postmanet is to provide the illusion of a *network* connection—these are very different abstractions involving different research agendas. The PR0 system leverages a single mobile storage device that always accompanies its user to transport storage system differences across multiple computers for a single user [14]. The Skunk and Segank systems [15, 17] aim to extend the PR0 system by dealing with the non-uniformity of connectivity technologies (including both short-range and long-range wireless alternatives) and allowing storage sharing among multiple users. The involvement of many mobile storage devices transported by the postal system, the network transport issues (such as those having to do with out-of-order delivery and retransmission), the provision of a network abstraction for multiple applications and multiple users, the simultaneous use of the postal system and the Internet, the peer-to-peer routing schemes, and the Postmanet applications are all distinct issues not seen in these earlier systems. In particular, we believe that the choice of providing a *network* abstraction as opposed to a *storage* abstraction is important: the network abstraction is at a sufficiently low level that may allow potentially greater degree of application flexibility, while an important goal of typical distributed storage systems is to entirely abstract away device or machine identities. The question of how to build a distributed storage system on top of the Postmanet, however, is still an interesting one.

C.7 Intellectual Merit and Broader Impacts

C.7.1 Relevance to the NSF "Digital Society and Technologies" (DST) Program

We believe that the Postmanet system proposed in this research addresses the NSF DST program focus in at least the following aspects. (We use the *slanted font* to quote the DST program description.)

• *Addressing barriers to information flow.* The DST program seeks to address *barriers to adoption and use* of information and knowledge *across geographic boundaries*, and to make *distributed resources* and data more *broadly available*. An important barrier is the lack of pervasive and high-quality bandwidth and, in the case of disadvantaged regions of the world, the lack of connectivity altogether. The proposed Postmanet has the potential of providing a form of inexpensive, pervasive, and plentiful bandwidth to the mass, potentially alleviating the "digital divide" phenomenon.

• *Understanding the science of collaborative work.* The DST program seeks to facilitate *distributed* and *collective* work among *humans, not just in one-to-one relationships, but also as part of teams, across geographic boundaries, in distant relationships.* A substantial portion of this proposal is devoted to the understanding and development of peer-to-peer systems and applications built on top of the Postmanet, both with and without infrastructural support. As an example "science" problem, the solutions to the peer-to-peer Postmanet routing problem rely heavily on many latest graph theoretic results (Section C.4.3).

• *Transforming applications.* The DST program seeks to understand and facilitate how the Internet transforms *the way the world communicates, works, and learns*, often in *distant relationships*. We have discussed in Section C.2.2 how

the inexpensive, pervasive, and plentiful bandwidth provided by the Postmanet may impact many existing and likely future applications. In particular, we have discussed in greater detail one of our target applications for the Postmanet, namely, the delivery of basic education in disadvantaged areas (Section C.5).

C.7.2 General NSF Concerns

• *Importance to advancing knowledge, and creative and original concepts.* See the above section (C.7.1).

• *Proposer qualifications.* We are collaborating with several organizations, both within and outside of Princeton, on some of the Postmanet applications, especially the distance learning applications. These include the McGraw Center for Enhancing Teaching and Learning, the Woodrow Wilson School of Public and International Affairs, and teacher training colleges in target areas.

• *Advancing discovery and understanding, promoting teaching and learning, and promoting participation of under-represented.* Some of the important applications that we target in this research seek to address needs of people in disadvantaged areas, both in the U.S. and in developing nations, where conventional high-quality connectivity is beyond people's reach. We believe that this focus can motivate students who are often less interested in traditional and more narrowly-focused research programs. We currently lose a number of graduate students each year (in both recruiting and retention) due to frustration with the lack of perceptible impact on real-world problems. Studies indicate that this has particular relevance for diversity issues including women and underrepresented minorities in engineering. We are planning a series of project-oriented undergraduate and graduate cross-disciplinary courses that will help develop and evaluate applications targeting disadvantaged areas. Some of these will be in collaboration with the Woodrow Wilson School, focusing on fusing technology and public policy.

• *Enhancing the infrastructure for research and education.* A key target application of the proposed effort is the delivery of basic education in disadvantaged areas (Section C.5).

• *Benefits to society.* Section C.2.2 describes the other applications that the proposed research can impact. In general, we believe that the possibilities enabled by an inexpensive communication channel that has a tremendous amount of bandwidth can be vast.

C.8 Results from Prior NSF Support

Randolph Wang is the PI of NSF grant CCR-0313089, titled: "ITR: Ubiquitous Mobile Storage." A goal of this effort is to build a coherent storage system on top of distributed mobile storage elements. A key challenge is to deal with the non-uniform wireless connectivity capabilities. Section C.6.2 briefly explains its relationship to the proposed effort. In NSF grant CCR-9984790, titled "CAREER: Low Latency I/O and Ubiquitous Storage," we built several novel high-performance disk arrays. Prototype implementations of these new array designs demonstrated significantly better performance than highly optimized conventional RAID systems.

Kai Li is the PI of NSF grant EIA-9975011, titled: "Adaptive, Performance Portable Software for Next-Generation and Immersive Applications." This project investigates scalable display systems, their driving applications, and adaptive load-balancing schemes for parallel rendering. We have developed software tools for a 24-projector scalable display wall system driven by a PC cluster. The software tools include features such as automatic screen alignment.

Margaret Martonosi is PI of NSF grant CCR-0205214, titled "ZebraNet: Position-aware Power-aware Wireless Computing for Wildlife Tracking." We have built hardware and software prototypes of GPS-based wildlife tracking sensors. We are targeting a January 2004 deployment at Princeton's Mpala Research Centre in Kenya. ZebraNet has also been quite successful in terms of involving undergraduates and under-represented minority groups in its work. In NSF grant CCR-0086031, titled "Designing Real-Power Systems: Static and Dynamic Techniques for Managing Power/Performance Tradeoffs," the core idea is to create systems that responsively trade off performance and power in ways that improve performance while keeping to the system's power budget. In NSF grant MIP-9708624, titled "Applications and Tools for Configurable Computing in Sequential and Parallel Computers," we focused on techniques for customizing hardware to the needs of particular software application or domain. NSF grant CCR-9502516, titled "CAREER: An Integrated Hardware and Software Performance Monitoring System," led to several key results in the area of performance measurement and monitoring. Most recently, we built the Wattch cycle-level power simulator that has become one of the most widely-used architecture-level power simulators in the research community.

D Bibliography

[1] ANDERSON, R. 'Trusted Computing' Frequently Asked Questions. http://www.cl.cam.ac.uk/~rja14/tcpa-faq.html, August 2003.

[2] CULLER, D. E., DUSSEAU, A., GOLDSTEIN, S. C., KRISHNAMURTHY, A., LUMETTA, S., VON EICKEN, T., AND YELICK, K. Parallel Programming in Split-C. In *Proc. of Supercomputing '93* (November 1993).

[3] FALL, K. A delay tolerant networking architecture for challenged internets. In *Proc. of ACM SIGCOMM 2003* (August 2003).

[4] GRAPH THEORY AND COMBINATORICS RESEARCH GROUP, APPLIED MATHEMATICS IV DEPARTMENT, UNIVERSITAT POLITCNICA DE CATALUNYA (UPC). The (Degree,Diameter) Problem for Vertex-symmetric Digraphs. http://www-mat.upc.es/grup_de_grafs/grafs/taula_vsd.html.

[5] HASSON, A. A., FLETCHER, R., AND PENTLAND, A. DakNet: A Road To Universal Broadband Connectivity. http://courses.media.mit.edu/2003fall/de/DakNet-Case.pdf, 2003.

[6] HORRIGAN, J. B. Adoption of Broadband to the Home. The Pew Internet and American Life Project, http://www.pewinternet.org/reports/toc.asp?Report=90, May 2003.

[7] JUANG, P., OKI, H., WANG, Y., MARTONOSI, M., PEH, L.-S., AND RUBENSTEIN, D. Energy-Efficient Computing for Wildlife Tracking: Design Tradeoffs and Early Experiences with ZebraNet. In *The Tenth International Conference on Architectural Support for Programming Languages and Operating Systems* (October 2002).

[8] KAASHOEK, M. F., AND KARGER, D. R. Koorde: A simple degree-optimal distributed hash table. In *Proc. of Intl. Workshop on Peer-to-Peer Systems* (2003).

[9] MADDEN, M. America's Online Pursuits: The changing picture of who's online and what they do. The Pew Internet and American Life Project, http://www.pewinternet.org/reports/toc.asp?Report=106, December 2003.

[10] PRAHALAD, C. K., AND HART, S. L. The Fortune at the Bottom of the Pyramid. *strategy+business issue 26* (2002).

[11] RATNASAMY, S., FRANCIS, P., HANDLEY, M., KARP, R., AND SHENKER, S. A scalable content-addressable network. In *Proc. ACM SIGCOMM 2001* (August 2001), pp. 161–172.

[12] ROWSTRON, A., AND DRUSCHEL, P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In *Proc. IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)* (November 2001), pp. 329–350.

[13] SHAH, R., ROY, S., JAIN, S., AND BRUNETTE, W. Data mules: Modeling a three-tier architecture for sparse sensor networks. In *IEEE SNPA Workshop* (May 2003).

[14] SOBTI, S., GARG, N., ZHANG, C., YU, X., KRISHNAMURTHY, A., AND WANG, R. Y. PersonalRAID: Mobile Storage for Distributed and Disconnected Computers. In *Proc. First Conference on File and Storage Technologies* (January 2002).

[15] SOBTI, S., GARG, N., ZHENG, F., LAI, J., SHAO, Y., ZHANG, C., ZISKIND, E., KRISHNAMURTHY, A., AND WANG, R. Segank: A Distributed Mobile Storage System. In *Proc. Third Conference on File and Storage Technologies* (March 2004).

[16] VON EICKEN, T., CULLER, D., GOLDSTEIN, S., AND SCHAUSER, K. E. Active Messages: A Mechanism for Integrated Communication and Computation. In *Proceedings of the Fifth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS-V)* (May 1992), pp. 256–266.

[17] WANG, R., GARG, N., SHAO, Y., ZISKIND, E., SOBTI, S., ZHENG, F., LAI, J., AND KRISHNAMURTHY, A. A Peer-to-Peer Mobile Storage System. In *Business Briefing: Data Management and Storage Technology 2002* (October 2002), World Markets Research Centre.

[18] WANG, R., LI, K., MARTONOSI, M., AND KRISHNAMURTHY, A. Distance Learning Technologies for Basic Education in Disadvantaged Areas. Tech. Rep. TR-685-03, Computer Science Department, Princeton University, November 2003.

[19] ZHAO, B. Y., HUANG, L., STRIBLING, J., RHEA, S. C., JOSEPH, A. D., AND KUBIATOWICZ, J. Tapestry: A resilient global-scale overlay for service deployment. *IEEE Journal on Selected Areas in Communications* (2004).