Turning the Postal System into a Generic Digital Communication Mechanism

Randolph Y. Wang* Sumeet Sobti* Nitin Garg* Elisha Ziskind* Junwen Lai*
Arvind Krishnamurthy†

ABSTRACT

The phenomenon that rural residents and people with low incomes lag behind in Internet access is known as the "digital divide." This problem is particularly acute in developing countries, where most of the world's population lives. Bridging this digital divide, especially by attempting to increase the accessibility of broadband connectivity, can be challenging. The improvement of wide-area connectivity is constrained by factors such as how quickly we can dig ditches to bury fibers in the ground; and the cost of furnishing "last-mile" wiring can be prohibitively high.

In this paper, we explore the use of digital storage media transported by the postal system as a general digital communication mechanism. While some companies have used the postal system to deliver software and movies, none of them has turned the postal system into a truly generic digital communication medium supporting a wide variety of applications. We call such a generic system a *Postmanet*. Compared to traditional wide-area connectivity options, the Postmanet has several important advantages, including wide global reach, great bandwidth potential and low cost.

Manually preparing mobile storage devices for shipment may appear deceptively simple, but with many applications, communicating parties and messages, manual management becomes infeasible, and systems support at several levels becomes necessary. We explore the simultaneous exploitation of the Internet and the Postmanet, so we can combine their latency and bandwidth advantages to enable sophisticated bandwidth-intensive applications.

Categories and Subject Descriptors

C.2.2 [**Computer Communication Networks**]: Network Protocols—*Applications, routing protocols*

*Department of Computer Science, Princeton University, {rywang, sobti, nitin, eziskind, lai}@cs.princeton.edu. Supported by NSF grants CCR-9984790 and CCR-0313089.

†Department of Computer Science, Yale University, arvind@cs.yale.edu. Supported by NSF grants CCR-9985304, ANI-0207399, and CCR-0209122.

SIGCOMM'04, Aug. 30–Sept. 3, 2004, Portland, Oregon, USA.

General Terms

Algorithms, Design, Performance

Keywords

Network Architecture, Peer-to-peer systems

1. INTRODUCTION

As the adoption rates of the Internet and broadband connections slow down in the U.S., latest studies [11, 7] suggest 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.

Bridging this digital divide, especially by attempting to increase the accessibility of broadband connectivity, can be 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. Satellite-based solutions have severe cost and aggregate bandwidth limitations.

In this paper, we explore the use of digital storage media (such as DVDs, flash memory devices, or 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 and Netflix) have used this approach to deliver software and movies on a large scale, and some researchers have reported shipping hard disks filled with astronomy data [5]—none of these existing attempts have turned 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.

• *Wide 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 and proven technology that works well today; and to use it for digital communication, one requires no significant new investment in exotic equipment.

• *Great 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 observing some fundamental technology trends. Storage density of flash memory and magnetic disks has been increasing at the annual rate of 60% to 100% for many years, and it is likely to continue in the foreseeable future. Besides flash memory and hard disks, even the next generation Blu-Ray DVDs can hold up to 27 GB per disc. Moreover, one can always ship multiple units at a time. The amount of information that can fit in a fixed amount of volume, or that can be shipped by the postal system for a fixed cost increases at an exponential rate, one that the realizable 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 expected 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 household sends (and receives) one 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 and Netflix highlights the low cost advantage of this approach. This cost advantage can be more apparent in certain foreign countries where dialup lines are charged based on time spent online and can be much more expensive. In addition to catering to "low end" users, the cost advantage of the postal system relative to that of a high-speed wide-area network also holds for corporate "power users" shipping large amounts of data [5].

• *Good scalability.* The postal system is highly decentralized, and it does not appear to easily suffer from potential bottlenecks. It 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 not only include rural residents, those who live in developing regions, and economically disadvantaged people, but also other groups of users, such as those who travel to places lacking connectivity for business or recreational purposes.

• *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 how 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 may 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.

Manually preparing mobile storage devices for shipment may appear deceptively simple, but with many applications, communicating parties and messages, manual management becomes infeasible, and systems support at several levels becomes necessary. In Section 2, we describe the user experience with the Postmanet, and some of its example applications. In Section 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 4, we examine options of "routing" data from the sender to the receiver in the Postmanet system. In particular, when there are a large number of communicating parties, minimizing the number of mobile storage devices sent or received at each site per postman visit can be an important consideration. This is an example of an unconventional routing metric that is unique to the Postmanet. We discuss how a large-scale peer-to-peer system (such as a file sharing system) can be built on top of the Postmanet. We shall see two recurring themes at these different levels of the system. One is the simultaneous exploitation 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, cost, and reliability. We describe related work in Section 5, and our conclusions in Section 6.

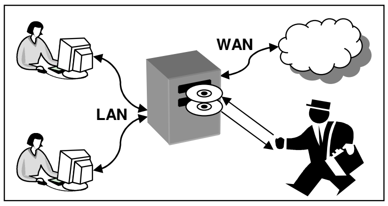

Figure 1: An imaginary Postmanet router box.

2. USAGE SCENARIOS AND APPLICATIONS

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 with one or more slots for inserting mobile storage media such as DVDs. (See Figure 1.) During the day, a user of the Postmanet router simply uses his applications in a way that is almost entirely oblivious of the presence of the Postmanet. At the end of the day, the Postmanet router box automatically ejects 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 device: the user's home computer may shoulder the task. We may use different types of mobile storage media: these may include read-only or read-write DVDs, hard disks of various form factors, 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.

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.)

• *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 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.* Large media files are excellent candidates for transmission over the HLHB Postmanet. The small messages generated by foreground file searches or background announcements of sites' contents can still use 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 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 multi-terabyte hard disks filled with hundreds or even thousands of movies being employed by the HLHB Postmanet. Large 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. For example, such a DRM contract may restrict 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, magazines and store catalogs (with richer presentation), software releases and updates, and public lectures given at universities. The possibilities enabled by an inexpensive communication channel with 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.

• *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.

In the above discussion of example applications, we have seen 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. 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.

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.

3. END POINT SUPPORT FOR TRANSPARENCY

At first glance, manual preparation of data for shipment on movable storage media may appear deceptively simple. But manually copying, naming and managing many messages, potentially for numerous applications and communicating peers, is cumbersome. The postal system represents a classic analogy of a datagram service: individual movable storage media may be damaged, lost, delayed, or delivered out of order. Human users or individual applications should not have to cope with these complications if they desire better guarantees and abstractions. What makes these issues even more complex is our desire to simultaneously exploit the Internet and to exploit the excess capacity of movable storage media to improve the latency, cost and reliability of the system. To cope with these complications, and to fully realize the potential of this approach, we need support at both systems and application levels. The type of sophistication we demand from this support is far greater than that can be provided by a local file system, which, for example, does not address any of the above transport issues. The lack of a programming interface makes it difficult for multiple applications to programatically and easily exploit, let alone to coordinate or share the use of, this communication mechanism.

3.1 Simultaneous Exploitation of the Internet

The Postmanet can be very valuable in absence of any traditional connectivity. With the aid of an LLLB connection such as a phone modem, however, the Postmanet becomes even more powerful and interesting. In the rest of this paper, we assume the simultaneous availability of such an LLLB link. 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. The question is how to use this scarce resource in an appropriate way.

When data arrives at a Postmanet receiver via the postal system, for example, the receiver should send an acknowledgement back to the sender over the Internet. This may further cause the sender to discard a local message copy that may have been saved for potential retransmission. More generally, the sender system may choose between the LLLB Internet and the HLHB Postmanet based on factors such as the amount of data 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: for example, 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.

3.2 Liberal Exploitation of Excess Capacity

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 (lower latency or better reliability).

One possible factor that may constrain the liberal copying of data onto movable storage media at the sender is available time. 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, 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 obsolete data that is sent earlier.

3.3 Handling Datagram Limitations

The limitations of postal system datagram delivery are exacerbated by our aggressive exploitation of the Internet and the excess capacity. 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 postal system, the receiver may need to discard the duplicates. Multiple storage media may have been delivered by the postal system out-of-order; and data delivered by the postal system and by the Internet may arrive out-of-order. Similar issues may occur at the application level also. For example, even in absence of duplicates or out-of-order delivery, the receiver application may discover that some of the newly arriving data is no longer needed due to application-specific reasons. In all these cases, the system must exercise care not to unnecessarily copy or use obsolete data. This is especially relevant due to our general approach of liberally "wasting" storage capacity in an effort to optimize for other metrics—this "wastage" needs to be checked by cleverly combining application-specific intelligence with Postmanet's transport-level algorithms.

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, not only are these problems further complicated by our aggressive exploitation of the Internet and the excess capacity, it is also the case that 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. For example, in some application scenarios where data coherence is relevant, the freshest data can be distributed over a number of devices: 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, if possible.

3.4 Other Issues

In addition to resolving transport-level issues, we also need to provide easy-to-use APIs. Programming models similar to existing asynchronous communication models [14] and the programming languages built on top of them [3] may be desirable. Under these models, handler codes associated with messages are asynchronously executed upon arrival of the messages to incorporate the newly arriving data into ongoing computations. Applications can be granted direct access to the data contained on the movable storage media to make data copying out of the mobile storage device potentially unnecessary. For read-only storage media, copy-on-write techniques may be necessary.

Security is another issue. The sender may need to compute fingerprints and/or encrypt data on outgoing mobile storage devices. The receiver may desire to ensure that (1) the incoming mobile storage device 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.

4. ROUTING

We have considered support at the communication end points in the last section. We now consider how data is routed from a sender to a receiver. The Postmanet has some unique routing metrics. For example, an important consideration is minimizing the number of movable storage media received or sent per site per postman visit.

4.1 Three Simple Routing Strategies

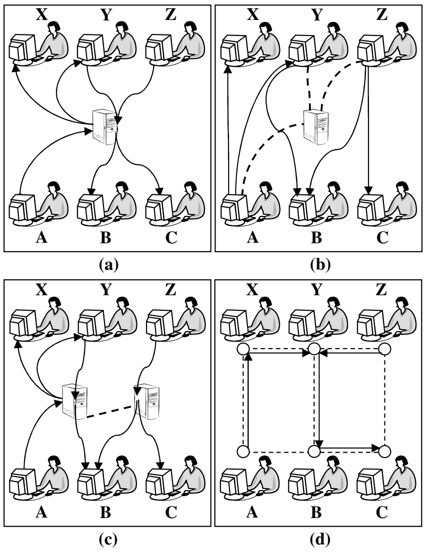

Figure 2: Routing strategies. A solid arrow denotes data communication carried by the Postmanet. A dash curve in (b) or (c) denotes routing information carried by the Internet. A dashed line between a pair of nodes in (d) denotes that it is permissible for these two nodes to receive movable storage media directly from each other. In all four panes, A sends different data items to X and Y, Y sends some other data to B, and Z sends different data items to 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. (d) Indirect peer-to-peer routing.

Let us consider the first three examples illustrated in Figure 2. Figure (a) is the centralized alternative. The server copies data from incoming storage devices to outgoing devices. The obvious disadvantages are unnecessary routing delays to and from a central server that can be located far away from the communicating parties, potential bottleneck effects developing at the central server, the extra cost incurred by the postal system, and the infrastructure cost of setting up and running the central server. This approach, however, 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 labels contain the postal address of the server. 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. Some of the end-to-end functionalities described earlier may also execute in the server. For example, the server may also enforce security policies so 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 sends two separate mobile storage devices directly to the intended recipients. (It is interesting to speculate how a peer-to-peer version of Netflix 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 data 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. 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 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.

4.2 Desired Routing Characteristics

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 substantial 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 site may need to send and receive many mobile storage devices per day.

An ideal Postmanet routing mechanism should possess the following characteristics: (1) it can accommodate a large number of simultaneous Postmanet communicators without requiring a site 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; (4) it does not burden Postmanet nodes in an unbalanced manner with data copying tasks that are beyond their own communication needs; and (5) it is robust when faced with misbehaving Postmanet end users. Some of these goals are unique to the Postmanet; these goals often conflict with each other; and we need to strike a proper balance among them.

4.3 Static Indirect Peer-to-Peer Routing

As illustrated in Figure 2(d), by requiring Postmanet 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 among the participating sites. This approach can eliminate the need for such an infrastructure, thus combining some of the advantages of the different strategies shown in Figures (a)-(c). In Figure (d), for example, Y sends to B a single disk, which contains both data from Z and data originating from Y. After B receives this disk and extracts data destined for B, it forwards a disk onto C, so C finally receives the data sent originally by Z.

Suppose the number of Postmanet nodes is N. In the following discussion, when we say a site "handles" k disks, we mean that the site may receive up to k storage devices and send up to k storage devices 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. In graph theoretic terms, the problem of simultaneously limiting the number of disks handled per node and maximum latency 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 latency, and the degree of a node corresponds to the number of disks it handles. For example, in a simple case, if N Postmanet nodes are organized into a k-dimensional directed mesh (or more precisely, a k-dimensional directed torus), so that each Postmanet node can only receive movable media from k of its immediate mesh neighbors and send to k of the remaining mesh neighbors, the worst latency is $k\sqrt[k]{N}$.

For a better solution, it is well known that with a constant node degree (or a constant number of disks handled per site), one can achieve $O(\log N)$ diameter (or maximum latency). Distributed Hash Table (DHT) topologies based on de Bruijn graphs can also probabilistically achieve the same latency with constant number of disks handled per node [10]. These DHT-based 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 may be 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 or topology repairs, while bulk data is communicated over the HLHB channels. Randomized constructions with constant degree for each node and $O(log\ N)$ diameter are also well known. A problem with the randomized approach is that it may be challenging to construct geography-aware routing topologies that can minimize unnecessary extra postal system delays and costs.

A potential complication facing any peer-to-peer system is coping with misbehaving participants. In a peer-to-peer Postmanet routing mechanism, where a node may fail to promptly forward data, replicating data on multiple outgoing devices along different routes can increase robustness. Protocols dealing with routing faults [2] may also be applied to such a Postmanet to isolate and penalize misbehaving nodes.

4.4 Dynamic Peer-to-Peer Routing

The possible routes described in Section 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, in the routing strategy shown in Figure 2(d), if C desires to send data to A, its data would normally be routed through B. But, there is no reason why C should not be allowed to send a disk *directly* to A if, on a given day, it does not overburden either of them. 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 is, again, a routing optimization problem unique to the Postmanet. 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 and end point management sections (Sections 2 and 3); 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.

5. RELATED WORK

Gray and his colleagues have shipped via the postal system entire NFS servers filled with terabytes of astronomy data [5]. NFS servers are chosen as mobile storage devices to minimize the amount of manual configuration a data recipient would need to perform. This is a goal that we share. Our interest is in generalizing these tailor-made solutions for specialized applications into a generic communication mechanism that can benefit many applications. A local file system interface that grants application access to the mobile storage devices may be inadequate: for example, tasks such as recipients' sending back acknowledgements over the Internet should be automated away by a transport-level system. We also note that the applicability of the Postmanet approach is by no means limited to data-intensive scientific applications: we have discussed a variety of applications that can be useful for average users, especially those who fall on the wrong side of the digital divide.

Rover is a toolkit for constructing applications targeting weak and intermittent wireless networks [8]. A key element of the system is an asynchronous communication mechanism that allows applications running on mobile wireless clients to continue to function as communication with a remote server occurs in the background. The need of an asynchronous communication mechanism applies to the high-latency Postmanet. The characteristics of the postal system, however, are different from those of a weak wireless network: the postal system provides a high-latency high-bandwidth datagram-like service. By simultaneously exploiting an available low-latency low-bandwidth Internet connection and the excess capacity of movable storage media, we can provide better higher-level services.

Recent efforts on "Delay-Tolerant Networks" (DTNs) [4, 6, 9, 12] 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. While "postal classes of service" have been mentioned, to the best of our knowledge, the postal system has so far only been mentioned as an *analogy*—no existing DTN that we are aware of literally uses the postal system. There are several important differences between existing DTNs and the Postmanet. First, 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 without needing investment in exotic equipment. Ad hoc routing, frequently a central focus of some DTNs, is not necessarily a top focus of the Postmanet. Instead, we are more concerned with somewhat less conventional routing metrics, such as the number of storage devices handled per site per postman visit.

Second, most existing DTNs are also frequently referred to as "challenged networks:" they may be limited by low bandwidth among mobile ad hoc elements, brief and/or intermittent contacts among these elements, small amounts of storage space on these nodes, and power consumption constraints. 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. Instead of carefully conserving resources such as storage space and bandwidth, we may in fact strive to "waste" some of these abundant resources in order to gain other advantages. Another unique aspect of the Postmanet is the possible availability of a complimentary low-latency low-bandwidth Internet connection: the techniques involved in the *parallel* exploitation of multiple connectivity technologies are different from those involved in the *sequential* forwarding of data from one connectivity technology to another.

The PersonalRAID system leverages a single mobile storage device that always accompanies its owner to transport storage system differences across multiple computers for a single user [13]. 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. 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.

6. CONCLUSION

In this paper, we have described how to turn storage media transported by the postal system into a generic high-bandwidth digital communication mechanism. The simultaneous exploitation of an available low-latency low-bandwidth Internet connection and the excess storage capacity allows us to improve the latency, cost and reliability of higher-level services. We have also described a range of routing alternatives that provide different tradeoffs of metrics that are unique to the Postmanet system. We believe the Postmanet can enable a variety of interesting bandwidth-intensive applications; and it presents an unconventional but promising approach to addressing the digital divide.

7. REFERENCES

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

[2] AVRAMOPOULOS, I., KOBAYASHI, H., WANG, R., AND KRISHNAMURTHY, A. Highly secure and efficient routing. *Proc. IEEE INFOCOM* (March 2004).

[3] 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).

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

[5] GRAY, J., AND PATTERSON, D. A Conversation with Jim Gray. *ACM Queue 1*, 4 (June 2003).

[6] 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.

[7] 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.

[8] JOSEPH, A. D., DELESPINASSE, A. F., TAUBER, J. A., GIFFORD, D. K., AND KAASHOEK, M. F. Rover: A Toolkit for Mobile Information Access. In *Proc. the 15th ACM Symposium on Operating Systems Principles* (December 1995), pp. 156–171.

[9] 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).

[10] 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).

[11] 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.

[12] 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).

[13] 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).

[14] 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.