Distance Learning Technologies for Basic Education in Disadvantaged Areas

Randolph Wang* Kai Li* Margaret Martonosi† Arvind Krishnamurthy‡

1 Introduction

Basic education plays a crucial role in uplifting disadvantaged areas from the grip of poverty. Traditional approaches centered around brick-and-mortar schools, however, may not easily meet the massive scalability needs of basic education in developing nations. In this document, we propose a digital technology-based distance learning project that serves to amplify the reach and productivity of teachers so that quality basic education can reach the disadvantaged children who have access to none today. We propose a collaborative effort that joins forces from Princeton University, Yale University, and several selected universities and organizations in China.

Our proposed effort utilizes an innovative blend of the "high-tech" and the "low-tech." The low-tech includes approaches such as using the transportation of mobile storage devices via the postal system to provide high-bandwidth interaction without relying on a well-developed networking infrastructure. The high-tech approaches seek to exploit an integration of multiple types of communication channels (including both data transmissions through the Internet and storage devices carried by the postal system) in a peer-to-peer system that bridges the space and time gap among geographically distributed teachers and their students. If successful, we believe that the proposed effort will not only dramatically improve the basic education landscape, but also put in place a digital communication infrastructure that will serve the needs of a wider array of applications in health care, commerce, information dissemination, and entertainment. This joint research effort should also provide a real-world training ground for cutting-edge computer science and pedagogical research. We hope that the proposed effort will serve as a model for interdisciplinary and international cross-border collaboration.

The rest of this proposal is structured as follows. Section 2 motivates our effort by briefly examining the important role played by basic education, the challenges faced by traditional approaches, and the promises of a distance-learning approach. Section 3 gives an overview of the main technical innovations. Section 4 gives a preliminary organizational plan, in terms of the initial participants that we plan to collaborate with, and a skeletal plan of execution. Section 5 gives a detailed description of some of the technical innovations related to computer science. (We have tried to keep the writing accessible to the less technically inclined readers, but readers should feel free to skip this section.) We close in Section 6 by addressing a number of frequently raised concerns and questions.

---

*Department of Computer Science, Princeton University, {rywang,li}@cs.princeton.edu.
†Department of Electrical Engineering, Princeton University, mrm@ee.princeton.edu.
‡Department of Computer Science, Yale University, arvind@cs.yale.edu.

2 Motivation

2.1 The Importance of Basic Education

In a modern society that is increasingly driven by the exploitation and development of science, technology, and information, there has never been a greater need than there is today for well trained and educated people. Basic education is the very foundation upon which a nation's future rests.

For developing nations, the role played by basic education is especially crucial. Peter Bell, the president of CARE, the renowned international organization dedicated to fighting poverty, cites improving access to basic education as one of the top three priorities if we are to end extreme poverty [1]. The lack of education is a crucial link in the chain that locks generation after generation in a vicious cycle that leads to the lack of job opportunities, poverty, and the inability to afford education for the next generations. If we are able to break this cycle by making basic education easily, widely, and cheaply available, we may be able to generate a self-sustaining "ripple-effect" that reaches beyond the current generation. Most experts agree that the best way of lifting depressed regions from poverty is to empower the local population by giving them the tools that they can use to take the initiative on their own and discover opportunities on their own. The investment made in improving basic education is consistent with this principle.

And there is more at stake than the impersonal calculus of economic development. Let us also not forget the human faces. Let us not forget the scenes of parents skipping meals or collecting empty plastic water bottles on streets, hoping to scrape together enough funds to send children to school. Extreme poverty is an assault on the dignity of human beings. Martin Luther King Jr. once said: "In a real sense, all life is interrelated. The agony of the poor impoverishes the rich; the betterment of the poor enriches the rich." All of us bear a responsibility in the struggle to end extreme poverty; and there are few tools that are more effective than improving access to basic education.

Furthermore, in an increasingly information-intensive society, many of those who are unable to receive basic education can feel a profound sense of disconnect from the society, a sense of disillusion and disenfranchisement, which in some cases may translate to sources of discontent and social instability. Access to basic education allows one to feel the sense of inclusiveness, the sense that one is a participating and contributing member of the society. This may provide a sense of spiritual well-being whose value may be more than what can be measured in dollar terms. An educated and informed citizenry is also necessary for improving the health of political systems. In a recent interview, for example, Chinese Premier Wen Jiabao cites the inadequate education level of the population as the top hindrance to more rapid progress in political reforms [14].

2.2 Challenges Faced by Traditional Approaches

A key difficulty faced by traditional approaches for improving basic education is "scalability." By some estimates [2], the population of the illiterate reaches nearly 200 million in China, and nearly 400 million in India. The condition of the female population is worse (more than 20% illiteracy rate in China and more than 50% in India). In all likelihood, these estimates are conservative. (The definition of "literacy" varies from place to place and can be very liberal.) There is also likely to be a sizable gap between literacy and a level of basic education that can open up reasonable job opportunities.

While there is no question that considerable strides have been made, especially in recent years, and there is no question that much more impressive progress should occur in the coming years, addressing the education delivery issue using traditional means is likely to continue to face serious difficulties. The building of traditional brick-and-mortar schools is costly and slow. The sheer number of the illiterate demands huge and continued resource commitments over many years. There is a lack of teachers, especially those that are well trained, well qualified, and highly motivated. There are many needy remote regions, which may have difficulty attracting and retaining good teachers.

Indeed, even developed nations are experiencing severe strain on educational resources. The U.S., for example, currently spends $880 billion per year on education; yet many believe that the U.S. basic education is inadequate. With far greater populations, far worse illiteracy rates, less per-capita resource, more remote and impoverished areas, poorer infrastructure, and vast regional disparity, it is likely to be a long struggle for developing nations such as China and India to catch up in the delivery of basic education.

2.3 Promises of Distance Learning

As networking and other digital technologies continue to advance, one possible solution to the problem of basic education dissemination is distance learning. It holds many promises. Content that is developed once can be used and reused at many places. Without the costs such as those needed for stationing many specialized teachers at many local schools, distance learning may be more cost-effective. It may also be more scalable if the content reaches a large audience. Today, children in some rural areas have to help parents with farm work, which, especially during certain times of the year, may interfere with going to school. Distance learning may allow these children to learn at a more flexible pace and schedule. For teachers who are enthusiastic about a career dedicated to helping needy children, but are reluctant to endure the hardship of living in less developed regions, distance teaching offers an attractive alternative. (This attraction may be especially relevant for regions that pose a greater hazard, places such as today's Afghanistan and Iraq.) The system also gives volunteers the option of more flexible time commitment. For example, a volunteer may donate one hour per day of her spare time to grade homework through the distance learning system. This is a much lighter commitment than what a volunteer would have to commit to today in a traditional system; so we may be able to attract more participants. For the teachers who are willing to travel to remote regions to serve the needy children today, distance learning also allows them to spend their time more productively by not wasting time on fruitless activities such as traveling.

It is important to note that the purpose of the digital distance learning mechanism is not to compete against or replace human teachers; on the contrary, the goal of the system is to amplify the reach and power of the limited number of qualified teachers that we do have. In other words, it is *not* our belief that the quality of a face-to-face session with a well-trained teacher can be exceeded by that of a lesson delivered over a distance learning channel—this is not our goal. Instead, for disadvantaged children who have no access to basic education at all today, access provided by distance learning should represent a significant improvement over a dismal baseline.

A positive byproduct of a digital distance learning undertaking, if well executed, is a software and hardware infrastructure that can serve other causes for the targeted disadvantaged areas. Some possible examples are: rudimentary health care (in the form of remote diagnosis), promotion of local commerce, and richer social interactions and entertainment. In general, it allows disadvantaged areas to be better connected to the larger world, allows local citizens to feel an enhanced sense of inclusiveness, and fosters innovative economic and social development. At the dawn of the industrial age, physical infrastructures in the form of railways, roads, and interstate highways served as powerful catalysts to bring progress to isolated parts of the world. In our current information age, digital infrastructures should serve equally critical roles. The development of such an infrastructure is also consistent with the principle of providing enablement instead of prescribing specific solutions—once the enablement technologies are in place, indigenous grassroots ingenuity may take over to exploit the infrastructure with more applications. Distance learning provides a compelling initial impetus for the development of such an infrastructure.

3 Technical Innovation Summary

3.1 Challenges of Distance Learning

Much of the existing distance learning effort has focused on the delivery of higher education in relatively resource-rich environments. There are at least two challenges if we were to successfully adopt distance learning as a means of providing basic education: (1) the need of providing sophisticated modes of interaction; and (2) the need of adapting to resource constraints, in particular, communication bandwidth constraints. To a large extent, these are conflicting goals, and existing approaches tend to favor one at the expense of the other.

There are two simple extreme approaches. In one extreme, content is disseminated via TV broadcasts or on storage media (such as CDs or DVDs) that can be delivered by the postal system. Content that is less bandwidth-intensive may be downloaded to learners' computers over the Internet. The amount of interaction afforded by this extreme is typically either non-existent or very limited. While such a self-guided approach may prove fruitful for mature learners who are self-motivated, self-disciplined, and savvy enough to explore and make the most out of a limited learning mechanism, 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 of the spectrum in terms of the amount of interaction, an existing approach to distance learning is similar to teleconferencing: a teacher and his students can engage in real-time video and audio interaction. The disadvantage of this approach is its consumption of a large amount of communication bandwidth—this cost and scalability handicap makes the approach unsuitable for resource-constrained environments in developing nations.

3.2 Relationship With Current Efforts

Under the auspices of the Chinese Education Ministry, several ambitious distance learning efforts were launched recently [5, 12, 21]. We believe that our proposed effort complements nicely with the existing endeavors. (1) While these current efforts focus on investing in the necessary enabling *hardware* technologies, our effort primarily focuses on the *software* aspect of the infrastructure, which, as experts have observed [12], is just as crucial (if not more). (2) As we have discussed in the previous section, the simplest deployment model based solely on playing DVD content lacks the element of student-teacher interaction, which we believe is crucial for a successful basic education process. The more advanced deployment model based on satellite networking, by itself, may also have its drawbacks: satellite bandwidth is limited and costly. Our proposal addresses interaction without requiring high-bandwidth interconnect. (3) While the current efforts seek to provide point-to-point interaction, our proposal seeks to unite all the participants, including schools, teachers, graders, and assistants into a virtual community, whose collective power to serve itself and to attract more participants may be far greater than the sum of many isolated sites. (4) One of our goals is to develop a generic communication infrastructure that can cater to other types of applications beyond distance learning.

As the lack of access to quality basic education is a world-wide problem, we believe that the proposed distance learning effort has the potential of developing interesting technologies that can catapult the various participating organizations in particular and China in general to leadership positions.

3.3 Overview of Technical Innovations

In this subsection, we provide an overview of the main technical innovations. We defer the more detailed descriptions of these technologies to Section 5. These technical approaches serve two high-level purposes. One is to provide meaningful interaction without depending on traditional high-bandwidth network communication. The other is to maximize the reach and productivity of the human teachers. Our first technique exploits alternative asynchronous high-bandwidth communication channels. Our second technique, a distributed peer-to-peer system for supporting interaction, not only allows more efficient information flow, but also allows increased flexibility and specialization in the teaching staff's time commitments. Our third technique, the exploitation of intelligence in user interface devices, may reduce both the bandwidth needs and manual intervention by remote teachers. In addition to serving the needs of distance learning, we believe these technologies can serve the basis of a more general digital communication infrastructure serving broader applications. The computer science research topics embodied in these technologies, including those in distributed systems and human interfaces, may also serve as graduate student research topics that provide a healthy blend of research and practice.

3.3.1 Providing Interactivity Using an Asynchronous Distributed Storage System

We ask the following series of questions. (1) Given a wide-area communication network that has relatively poor bandwidth, what modes of interaction can we support that can provide valuable assistance to a basic education process? (2) Suppose we generalize the definition of "communication networks" to include storage media shipped in a postal system, what modes of interaction can we support? How is this new form of communication integrated into the rest of the system, which includes the use of traditional networks and local storage elements? (3) How do we generalize this infrastructure so that it may benefit applications beyond distance learning?

We shall refer to bytes delivered by mechanisms such as a postal system as a High Latency High Bandwidth (HLHB) channel. (As we explain later in Section 5.1.1, there are other examples of HLHB channels than the postal system.) Conversely, we shall refer to bytes delivered by mechanisms such as a traditional Internet as a Low Latency Low Bandwidth (LLLB) channel. At a first glance, the use of a so-called HLHB channel may appear to be a stopgap solution in a resource-poor developing region. However, we do not believe this to be the case. The appeals of these HLHB channels are based on fundamental technology trends of the growth rates of storage density and wide area bandwidth: the former is far greater than the latter. The implication is that these HLHB channels are highly attractive for developed and developing regions alike, and they should become even more attractive as the gap between storage technology and wide-area networking technology widens in the future. Our aim is to develop an *asynchronous distributed storage system* to exploit a seamless integration of all these communication channels and to develop interesting applications on top of them. Distance learning is only one example application; it is our hope that the asynchronous distributed storage infrastructure developed by this project will enable many other commerce, informational, and entertainment applications to follow.

3.3.2 Peer-to-Peer Interactions

As discussed earlier, our goal is not to replace humans. Rather, our goal is to amplify the effectiveness of the limited human resources that we do have. We would like to attract geographically distributed participants that include not only students at various levels and teachers of various specialties, but also volunteers and staff who may serve as graders and teaching assistants. 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 system to act like a "marketplace" where the demand and supply of various services (such as the grading of homeworks) are being matched up. In a sense, this is analogous to existing commercial marketplaces such as auction sites.[1]

A naive approach to the construction of such a marketplace is to force interactions through a centralized server so, for example, a student would submit homework (digitized via scanning, for example) to a central server using any of the communication mechanisms discussed above (Section 3.3.1). The centralized authority would delegate the grading job to a grader who is potentially geographically located at yet a third location so the homework data is re-routed to this third location. (The graded homework, which includes various levels of feedback, may also need to be returned to the student.) The inefficiency involved in this naive central server-based approach is obvious.

A better approach is, obviously, to "connect" the student and his grader directly. Such "connections," however, need to be coordinated. In technical terms, what we need is a peer-to-peer storage system, a peer-to-peer routing mechanism, and a peer-to-peer interactive application built on top of them. We face at least two challenges that are not well-addressed in existing systems: (1) addressing application-specific semantics (such as those related to homework grading) in peer-to-peer routing decisions; and (2) the interaction of a peer-to-peer architecture and the use of asynchronous communication channels (as discussed in Section 3.3.1).

We believe that by building the distance learning system around a peer-to-peer architecture, we may gain a number of desirable features in the resulting system. It allows broad and flexible participation. The peer-to-peer architecture presents a framework which people who desire to play various roles can "plug" themselves into. The system grows and scales in a decentralized manner as the number of participants increases. These participants may be widely distributed in both space and time. The goal of the distance learning system is to bridge this time and space gap.

Some may develop new teaching materials and make it available for potential learners. Some may start new local schools to tap into an online feed. Some students may choose to join as

---

[1] 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 some others who may collect a small fee in exchange for their time. The kids, naturally, should get their education for free.

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. Some of the service providers may be volunteers, while others may draw a modest payment for their efforts from funds provided by governments or other non-profit organizations. Indeed, we hope that some professional teachers may choose to teach via the distance learning mechanism for a living. In addition to student-educator interactions, the system also makes it possible for student-student and educator-educator interactions: for example, students in remote areas may interact, both academically and socially, with students in other areas.

Such a system can also encourage specialization, a principle in modern economies that promotes increased efficiency. Traditionally, teachers tend to shoulder many distinct duties: development of teaching materials, delivery of lectures, grading of homeworks, and holding office hours, to name a few. 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 in our target environments. For example, the grading of simple forms of homework could be handled by individuals with relatively lesser skills or less experience. Well-trained and experienced teachers, who are good at interacting with students, 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.

Of course, as we have mentioned earlier, the participating teaching staff should be subject to a process of careful screening and quality control. We could use the same peer-to-peer infrastructure to train a local staff who may interact with children directly (face-to-face). The role of a local staff is important: while the standardized content of a rigorous curriculum may be delivered remotely (via one or more of the possible communication channels discussed in Section 3.3.1), a local staff, who may not necessarily be fluent in the details of the content, will handle functions such as policing the children, managing equipment, and interacting with children about many administrative details. For more knowledgeable staff, their roles can be expanded. One of the possible sources for the recruitment of the local staff is the very children that we teach—some of the upperclassmen, for example, may serve as teaching assistants for their junior counterparts. The goal is to create a self-sustaining and self-amplifying cycle to address one of our most important resource concerns: the human resources.

What we have described above is a wide variety of real-world peer-to-peer interactions that should occur in a large-scale distance learning system. These interactions should be naturally reflected in a system possessing a peer-to-peer architecture. When completed, we believe that such an infrastructure may enable other types of peer-to-peer interactions, interactions in commerce, informational, and entertainment applications.

3.3.3 Intelligent Human Interface Devices

Yet another way to provide rich interactions without consuming excessive wide-area bandwidth is to utilize intelligent human interface devices. Consider an example in which a child draws a character on the board and a teacher needs to provide feedback. In a simple solution, one may point an inexpensive video camera at the board and transmit the evolving board content to a remote teacher. This solution demands a large amount of communication bandwidth. There are three approaches for adding intelligence to interface devices to reduce their bandwidth requirements without sacrificing good interactivity.

The first is to capture the image data using a better (more compact) data representation. An example is the "digital ink" technology employed in Microsoft TabletPCs. There are, however, several issues with TabletPCs that we must address. The first is that we need an interface device that is more similar to a blackboard, a large form factor device that is sharable by a group. This is desirable because the group can collectively learn and benefit from what each member is doing. The solution must also be inexpensive: a solution that requires equipping each student with a TabletPC, for example, is likely to be beyond a budget that will be available to us, and is not necessarily the best way of spending money even if we could afford it. Other desirable features include the ability to convey dynamic content evolution (as opposed to static snapshots) and the ability to adapt to various communication channel capabilities (including those of asynchronous channels).

The second approach is to provide some form of automatic content recognition and rudimentary local feedback without requiring constant feedback from a remote teacher. Unlike traditional handwriting or speech recognition, whose goal is to deduce a *conclusion*, our goal is to recognize *patterns* in a learning process. To this end, we need to take into account *how* actions are performed, in addition to *what* results are produced. Machine learning techniques that our colleagues have employed in recent content recognition projects [7] may play an important role. Of course, we need to carefully determine what interactivity can be provided locally, what must be done remotely, and how to integrate the two. We should also exploit what can be learned from one group of students and potentially leverage that experience in a different (or later) group. For example, we may attempt to anticipate common learning mistakes and provide corresponding prepared feedbacks. Another factor that we must carefully consider is how to take advantage of help that a local staff can provide—such a staff may lack the domain expertise of particular subjects, but they possess "common sense knowledge" and pattern recognition abilities that have been considered challenging for artificial intelligence efforts.

The third approach is to use *digital avatars*, virtual human-like beings that, for example, utter words and perform actions, emulating real-world remote teachers. Transmitting the control information for these avatars can be much more efficient than sending live video feeds of remote teachers. At the same time, they provide more personable and interesting interactivity to children than simpler means such as disembodied voices.

3.3.4 Other Areas of Interest

There are a few other areas that we plan to investigate. One such area concerns how to provide or improve physical networking connectivity to remote regions. The other is to provide an authoring environment for the creation of distance learning teaching materials, including interactive content that can exploit various forms of asynchronous communication mechanisms. The content authoring process may utilize a number of ways of observing and capturing regular teaching sessions: (1) passive and non-intrusive recording; (2) "cooperative recording," which, for example, may require teachers to use some of the intelligent interface devices (discussed above) to aid the capture process; and (3) "post production augmenting," during which additional content is created after a regular teaching session is completed. Multiple versions of the content, tailored for various bandwidth conditions, or transcoded in a content-aware fashion, may be created during the authoring process. In addition to aiding distance learning, this digital content can be useful aids for a traditional teaching process as well. Just as is the case with the delivery system, a challenge for the authoring system is to make at least a substantial portion of it content-neutral and culture-neutral so that it can be easily re-targeted for a different curriculum. In these areas, if possible, we would like to leverage as much as possible some of the latest technologies—although it is not yet clear what our own innovations will be, we expect to have a clearer research agenda in these areas after we gain some initial experience.

4 Organizational Plans

4.1 Participants

We would like to invite and bring together the following groups of participants for the initial phases of the project.

• *Target schools.* These are the schools that we would like to set up at places requiring assistance. They could be in cities, or more likely, in remote rural areas. Participants from these local sites would play an important role in helping us understand the local conditions and needs.

• *Teachers colleges.* These are the colleges where the next generations of teachers receive their training. Examples include *East China Normal University* in Shanghai and *Beijing Normal University* in Beijing. By participating and helping in a new and exciting technology-driven distance learning effort, selected (volunteering) students from these universities may gain valuable experience in the following areas: (1) developing digital teaching materials tailored for the distance learning channels; (2) practicing teaching with real-world children, via the distance learning mechanisms; and (3) familiarizing themselves with the use of digital technologies. We note that these are not necessarily entirely novel responsibilities for these students—to a large degree, these tasks, especially the aspect of practicing teaching in real-world schools, are already integral parts of well-established routine training. The introduction of a new distance learning aspect merely adds a new dimension to their current training; in other words, the role played by technology should be complementary, not substitutive. We also hope that the opportunity of experimenting with cutting-edge technology may present an added incentive that would encourage students to pursue an education at these teachers college programs.

• *School teachers.* We would also like to invite participation by experienced school teachers. Their primary role would be to help develop teaching materials. Our ultimate goal is to allow children in remote regions to enjoy the best educational experience delivered by the best teachers, so the participation of good teachers is important: we would like to study how the pedagogical approaches of the best teachers can be codified in a distance learning curriculum. We also hope that some of these teachers may actually choose to teach via the distance learning channels.

• *Technical universities.* These would include computer science departments of some of the top universities in both China and the U.S., such as *Tsinghua University* in Beijing, *Jiaotong University* in Shanghai, and *Princeton University* and *Yale University* in the U.S. As we have explained in Section 3, to meet the challenges of providing sophisticated interaction while facing wide-area network bandwidth constraints, we plan to explore a number of technical innovations in distributed systems and human interface research. The responsibility of the participating graduate students and faculty in these computer science departments is to work on these technical areas while closely collaborating with the distance learning teaching staff. Research on these technical topics should also lead to publication of research papers and theses for advanced degrees.

We believe that all these groups can benefit a great deal by learning from each other. The teachers college students and teachers gain exposure to technology. The computer science students and faculty are provided a real-world grounding for their technical research. Children at target schools benefit from the expertise of experienced teachers and the enthusiasm of college youths. And by reaching out to those beyond our familiar surroundings, all of the distance learning teaching staff may gain a deeper appreciation of the resourcefulness of those whom we "meet" and hope to help, and have the satisfaction of knowing that we have perhaps made a small difference in those lives.

4.2 Steps of Execution

- *Planning phase.* We plan to talk to and meet with a variety of organizations to form an initial development team. This team will have members drawn from each of the groups described above. This team will select a small number of areas where target schools will be set up as initial test cases. The team will visit these areas to more accurately assess the local needs and evaluate the resources available. (The resources that we need in these initial phases are mainly in the form of the people who will spend time on the project; we believe that the equipment needs in these initial phases will be modest.) The team will refine the design of what features that the successive versions of the proposed distance learning system should contain. The ideas contained in this proposal should be treated merely as seeds for further discussion: the team members may come up with other ideas, prioritize some of the existing ideas, and forgo others.

- *Development phase.* The tasks of this phase are: (1) developing teaching materials suitable for the digital distance learning channels; (2) developing technologies for facilitating interactions; (3) conducting small-scale test teaching sessions to assess the strengths and weaknesses of the system; and (4) refining plans for improvements and iterating. It is important that we keep the initial incarnations of the system simple: we would like to use the lessons learned from the real-world test teaching sessions to drive each iterative refinement of the system.

- *Deployment phase.* Partially overlapping in time with the development phase is a number of preparatory tasks that we will be working on to gradually transition into the deployment phase. One of these is identifying an expanding circle of target schools that will be taught via the distance learning channels. The other is seeking funding for equipment and expanded teaching personnel. While it is not a serious concern for us if the content authoring stage requires relatively expensive equipment, one of the design constraints that we will force ourselves to observe is minimizing the cost of the equipment deployed at each remote school, because this is the part of the cost that would be multiplied many times as we seek increasingly wider deployment. We expect that equipment donations from various sources would largely meet our computing needs. (Hardware costs are dropping very rapidly.) Research and development work may also continue during the deployment phase: issues such as scalability will become increasingly important as the system grows.

5 Technical Details

In Section 3.3, we have given an overview of the technical innovations. In this section, we give a more detailed description of each of these areas.

5.1 Providing Interactivity Using an Asynchronous Distributed Storage System

Let us consider two communication channels: one is effected by shipping storage media (such as DVDS or even hard disks) via the postal system; the other is a wide-area Internet connection. If we compare the number of bytes that can be shipped via each of these two channels 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 believe, however, that this not to be the case. Our belief stems from an observation of some fundamental technology trends. Magnetic 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 be 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. On the other hand, 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, factors that doom the wide-area bandwidth growth to far slower rates. Also, the cost of furnishing "last-mile" wiring can be prohibitively high, and the progress has been excruciatingly slow. 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.

A communication mechanism such as shipping storage media via the postal system can yield enormous bandwidths, but 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. And we call a traditional Internet connection a Low Latency Low Bandwidth (LLLB) channel. In addition to the higher *per channel* bandwidth advantage of an HLHB channel, it has some other advantages. One is higher *aggregate* bandwidth: while the aggregate bandwidth of the Internet is limited by factors such as backbone bandwidth, the HLHB channels can be much more decentralized and can scale to greater aggregate bandwidth by adding channels. Another potential advantage of an HLHB channel is its lower cost. The exploitation of these HLHB channels is particularly important for remote or developing regions, where the development of a high quality wired infrastructure may be years away, although as we explained earlier, this approach is also highly relevant for developed regions.

Using HLHB channels to send digital content is not a new idea. For some time, companies (such as AOL.com and netflix.com) have used the postal system to deliver software and movies on a large scale. The questions that have not been examined, which we plan to address, are: (1) how do we integrate the HLHB and LLLB channels, along with local storage, to form a cohesive distributed storage system? and (2) how do we provide interactive instruction?

5.1.1 "Downlinks"

"Downlinks" refer to the delivery of teaching material from a remote teacher to children. (We will consider "uplinks" and interactivity in Section 5.1.4.) Let us consider the example illustrated in Figure 1. Suppose a class that has been offered in a previous semester is being taught again.

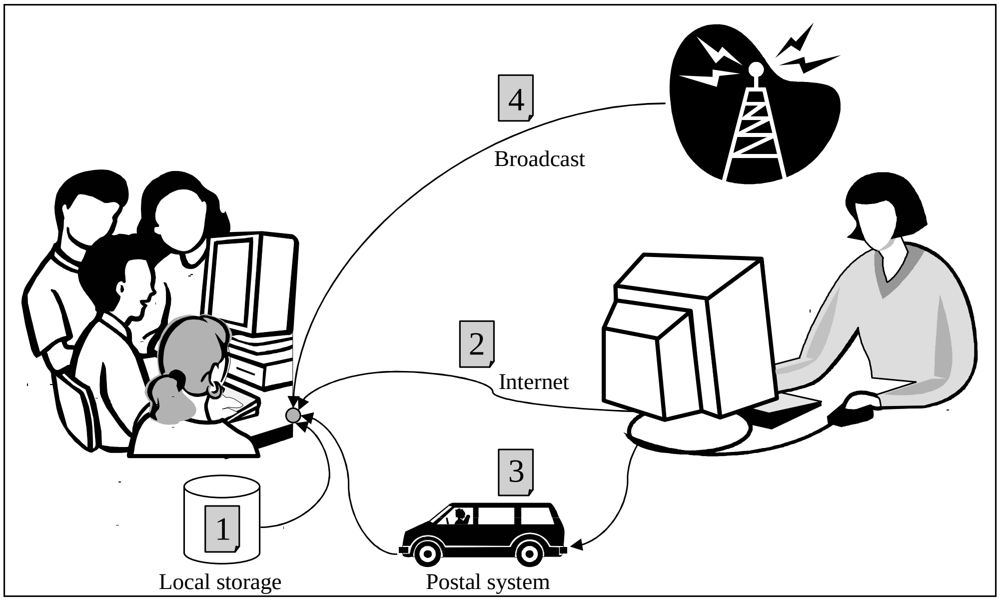


Figure 1: The integration of multiple downlink communication channels.

Because it has been taught before, much of the course material is already stored on a local disk at the school (illustrated by data block 1 in Figure 1).

The teacher may determine that a two-minute segment of an existing lesson needs to be replaced, either because it is incorrect, or because it needs to be tailored to the current situation. These are the kind of adjustments that good teachers would want to do all the time and are an integral part of the art of a creative teaching process. This two-minute segment might be small enough that we may choose to deliver it over the Internet (illustrated by data block 2 in the figure). The transmission would be initiated early enough to ensure that it is completed before the children receive the lesson containing the segment.

The teacher may choose to deliver a new two-hour segment that contains either new content or feedback for children's homework. Due to the large size of the data, it is placed on a mobile storage media and shipped by the postal system (illustrated by data block 3 in the figure). There may be other possible types of downlink channels. Figure 1 shows yet another possible channel, a radio or TV broadcast channel, which may transmit content either in analog or encoded digital form (illustrated by data block 4 in the figure). Also, the mobile storage devices need not be exclusively carried by the postal system: the devices may be carried and shared by anyone, and one of the goals of our system is to "knit" all these distributed devices into a coherent whole [16, 17, 20].

5.1.2 Complexities of Manual Management

At a first glance, the simultaneous use of these multiple channels may seem straightforward enough for a teaching staff to manage manually. The reality can be much more complicated.

If we were to naively encode an entire lesson in a single large video file, for example, it would be unwieldy to identify and replace a two-minute segment. So we would need to have a large number of fine-grained data elements that can be reassembled, rearranged, or replaced in flexible ways. Being able to manually track these data items is not easy. When a teacher authors a number of new data snippets, in a naive manual solution for example, she would have to assign some meaningful names to these snippets. She would have to think about which of the multiple channels to pick to ship the data. She would have to prepare instructions for the person who receives these snippets. She would have to manually copy data onto mobile storage media for shipment. She would have to worry about whether and when the content has reached the destination. She would have to respond to the loss of data in transit.

When the data arrives, a local staff would have to decipher the instructions on what to do with it. She would have to arrange for acknowledgements to be sent back. She would manually copy data out of the arriving mobile storage media. She would have to worry about out-of-order delivery of multiple mobile storage devices by the postal system. She would have to worry about out-of-order delivery of the Internet data with respect to data delivered by the postal system. She would have to worry about answering queries from the remote teacher who may want to know whether certain data has arrived. Finally, when both the local staff and the remote teacher are satisfied that all the pieces are properly put in place, they may schedule the delivery of a lesson. During the delivery, more adjustments may need to be made and more new data may need to be transmitted, so the two sides may need to perform more manual work to accommodate these adjustments. It is clear that the need for all these manual interventions might severely limit the benefit that we can derive from these multiple communication channels.

5.1.3 An Asynchronous Distributed Storage System for Multiple Communication Channels

A goal of the storage system that we propose to build is to automate away almost all of the manual tasks described above. Using this system, a remote teacher would simply author a new segment of an existing lesson. She does not have to worry about naming the new segments. She may provide some hints to the system such as at what time in the future she expects the new segment to be used by remote pupils.

The teacher does not have to worry about which communication channels to use to meet her needs; the system would take care of the choices depending on how much data is involved and how much time is available. Indeed, the system may choose to use multiple channels in parallel: it may prepare a low resolution version to be shipped over an LLLB channel, while a high resolution version is shipped simultaneously over an HLHB channel. The two versions would "race" against each other to play a tradeoff between quality and availability. If the Internet is "max'ed out," the system needs to carefully prioritize what is sent through the Internet connection and at what level of resolution. One way of looking at this problem is to view the Internet connection as a "cache" of the mailman connection. The system may choose to ship multiple copies of the data (spaced out in time) via the postal system to provide added reliability.

The teacher does not have to manually initiate any communication. If the Internet is to be used, the system starts network communication automatically. If the postal system is to be used, the teacher does not have to remember what data to copy to a mobile storage media, and she does not have to perform the copying manually. At the end of her working day, for example, the system would automatically prepare a mobile disk that contains the new content. If there are multiple teachers who are authoring new content, the system would automatically gather the new content from these multiple teachers onto a single mobile storage media (using a high-speed local area network, for example).

At the end of the day, a mailman would make a routinely scheduled stop to pick up the mobile storage media for delivery. Indeed, the copying does not have to start at the end of the day—continuous background copying could have occurred throughout the day, so that we avoid the time-consuming burst of copying toward the end of the day before the mailman arrives. We also note that due to the relatively large capacity of storage devices, as long as there is enough time to perform the copying, the system should be very liberal in deciding what gets copied. For example, even when there is only a very small chance that a piece of data will eventually be used by the pupils at a remote school, it does not hurt to go ahead and copy the data anyhow. It is perhaps ironic that we are trying to achieve lower latency by using an HLHB channel—lower latency because it will be faster for the pupil to access data on the mobile storage media than it is over the Internet (as long as the asynchronous postal system latency is masked). An important principle of this storage system is to exploit the plentiful resources (storage capacity) to circumvent the limitation imposed by scarce resources (wide-area bandwidth).

Once transmission is initiated in one or more communication channels, the teacher does not have to worry about monitoring progress. The postal system could provide tracking that aids the system's automatic progress monitoring. If the sending side fails to receive an acknowledgement (which should arrive over the Internet) after a certain period of time, it may take one of several possible actions. If a mobile storage media is lost by the postal system, and there is enough time left before the data is to be used by the receiver, the system can simply "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 mailman visit. If a postal mobile disk is lost, and there is not enough time to retransmit in the postal system, the system may choose to send a much lower-resolution version via the Internet. If all possible remedies are exhausted, the teacher would receive a notification. The teacher may choose one of several alternatives: reschedule the date of the lesson in question so more time for retransmission is made available, or skip the use of the new segment and go ahead with the lesson based on the old content.

Regardless of which channel(s) are used, once the data arrives, there should be minimum manual intervention by the local staff. This should be trivially true for data that arrives over the Internet, which naturally should trigger a receipt acknowledgement to be sent over the Internet. If the data arrives in a mobile storage media delivered by the postal system, all that the local staff needs to do is to have the newly arriving storage device "hooked up." If the media is a DVD, inserting it into a drive should suffice. If the media is a Microdrive-like small disk, the local staff may need to be trained to use an adapter to have it plugged in. In all cases, the "connection" of a media should automatically trigger the execution of some code that would perform a number of tasks. One of these is to send an acknowledgement to the sender (naturally, over the Internet). (The remote teacher may now know that her scheduled event is ready to go.) Another triggered task may be to copy data from the mobile storage media into a local store maintained at the school. Another type of triggered tasks may be, for example, printing out the graded homeworks for dispersal, or the automatic scheduling of a review session to be run the next day.

Also note that it is not necessary to copy the data from the mobile storage media into the local store before it can be used—if a scheduled teaching session based on the newly arriving data is to be run shortly and if there is not enough time for copying, the data can be accessed directly from the mobile storage media. Indeed, the remote teacher may have changed some of the content again after she sent off the mobile disk. And this freshest content may have arrived over the Internet before the arrival of the mobile storage media, which now contains some stale data. Or alternatively, some of the freshest content may have arrived earlier in yet another mobile storage media, delivered by the postal system *out of order*. The system must exercise care not to use or waste time copying the stale data. Or alternatively, some of the freshest content may still be residing on the remote teacher's computer. Or alternatively, a newly arriving storage media may contain duplicate data (due to either proactive replication or premature retransmission).

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 our context of leveraging the transportation of mobile storage devices, however, the boundary between storage and networks is blurred. The latency and the amount of data involved in a "packet" 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. The freshest data can be distributed over a number of devices: the local school store, an in-transit "data distribution center" (see Section 5.2), the newly arriving mobile storage media, and even the remote teacher's computer. When a lesson based on this distributed data is delivered, the system needs 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.

Another issue the system needs to address is security. For example, when content arrives in a mobile disk delivered by the postal system, we should have the confidence that (1) the data is from someone whom we are willing to receive data from; (2) the senders cannot forge identities; and (3) the data has not been tampered with. If desired, a generic version of this system may be more liberal in its policy of accepting data so, for example, "spam" could be allowed but given a lower priority of receiver processing. Indiscriminate spam is also less likely to be a problem when the sender must pay the postal system to get his message across.

The communication discussed above is mainly push-based; it is possible to provide other modes of communication using these multiple channels. For example, consider the implementation of asynchronous "reads" or fetches. A read request is sent over the Internet (or some other channel). Upon arrival, if the desired data is large, the read request would automatically trigger the desired data to be staged on a mobile storage device, which would be transported back by the postal system. 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 [18] and the programming languages built on top of them [4]. Under this programming model, handler codes attached to messages (called "Active 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.

Note that although we have phrased the discussion of the proposed storage system in the context of distance learning, what we are proposing is actually a general storage system that can cater to a wider array of applications. It shares some similarities with some of the disconnected mobile storage systems that we have recently built [9, 16, 17, 20]. We expect the proposed asynchronous distributed storage system to be able to support, more generally than these previous systems, a diverse set of bandwidth-intensive publish/subscribe applications.

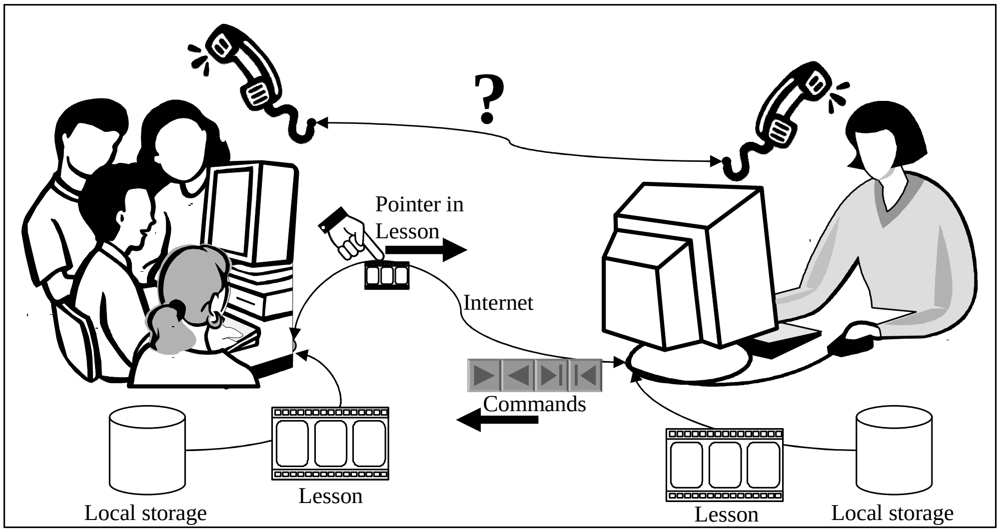


Figure 2: Synchronous interaction.

5.1.4 "Uplinks" and Interactivity

So far, we have focused on the downlink direction, the direction of communications from the teachers to the pupils. We now consider the uplink direction, and how teachers may provide feedback to pupils using asynchronous communication channels.

Some of the interactions may have relatively low bandwidth demands, and they may occur in real time. For example, voice carried by phone or the Internet may allow pupils to interact with their remote teacher in real time. The local staff may play a filtering role so that only selected questions are forwarded to the remote teacher. 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 2, if a pupil asks a specific question about a subject matter in the middle of the content stream that is being played off a local disk, 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 have access to the same data stored on their respective local disks, all the teacher needs are "pointers" into the content stream; these pointers are transmitted from the pupils along with their questions in real time. In addition to providing voice feedback, the teacher can also send back computer commands over the Internet to control the "playback" of the lesson at the students' site. These commands may rearrange the order of the snippets, or even play a pre-canned response to an anticipated question.

With limited communication bandwidth, 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 3. In this figure, there are two groups of students: group A and group B. (These two groups may or may not be co-located.) Group A

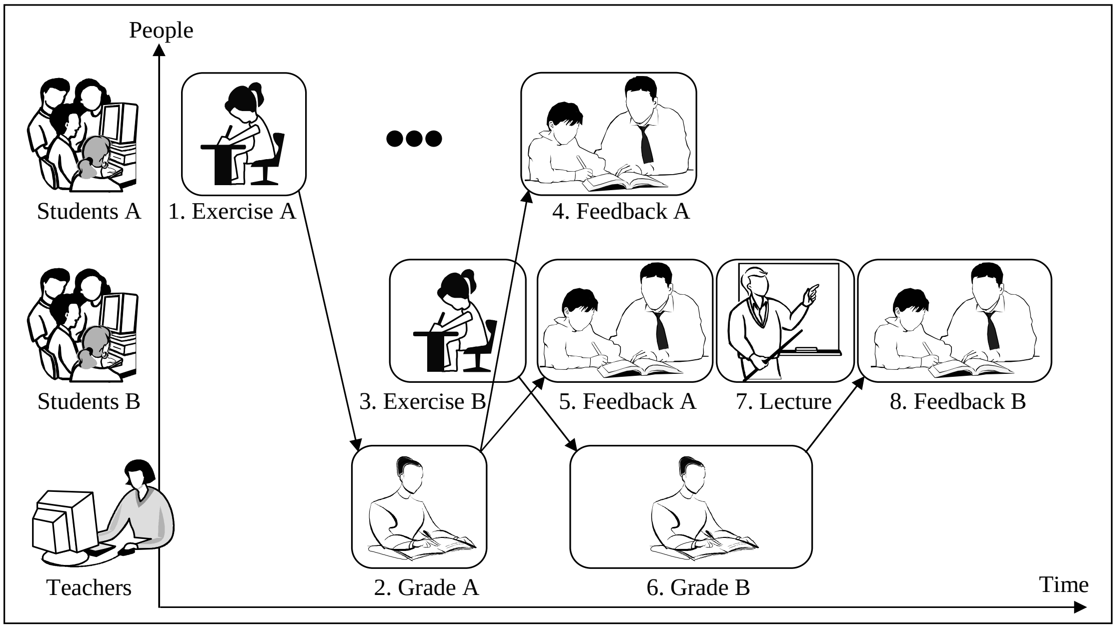
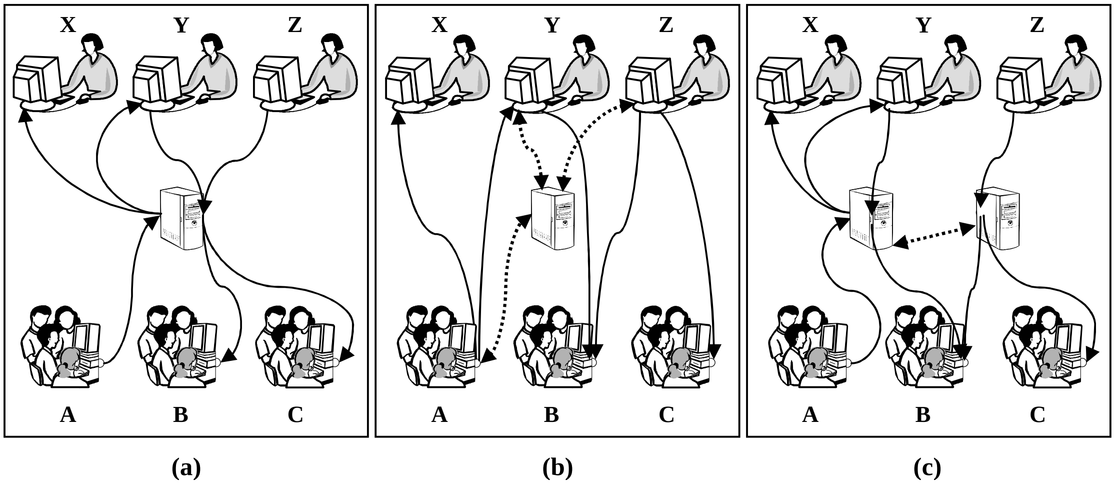


Figure 3: An example Gantt chart of providing interactivity using asynchronous communication channels. The arrows denote communication. The numbers denote the order of events.

performs an exercise (1), which, when finished, is transmitted to the remote teacher for grading and feedback preparation (2). (The asynchronous communication channel may be any one of those discussed in Section 5.1.3, and 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 that is used.)

In the mean time, group B starts the same exercise (3). Before group B finishes the exercise, the remote teacher has finished grading the exercise submitted by group A, and the teacher's feedback for group A is transmitted back to both groups A and B (4 and 5). As soon as group B finishes their exercise, the system plays the feedback for group A in front of group B (5). From the point of view of a child in group B, the interaction that he observes between other children and the teacher should be similar to what he would see in a conventional classroom.

In the mean time, the exercise performed by group B is transmitted to the teacher, and the teacher grades it (6). The students in group B may receive more (pre-delivered) lectures (7) before they review the feedback that the remote teacher has prepared specifically for them (8). For a child in group B, he sees an uninterrupted stream of carefully prepared material. The stream includes instantaneous general feedback (although the feedback consists of the teacher's interactions with other kids) 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. The management of communication in both directions should be automated in the manner described in Section 5.1.3.



(a) (b) (c)

Figure 4: Alternative routing architectures. A solid arrow denotes data communication that is carried either by the Internet or other asynchronous communication channels, such as a mobile storage device carried by the postal system. A dotted arrow denotes routing information (which is most likely carried by the Internet). (a) Centralized routing via a single data distribution center. (b) Peer-to-peer routing. (c) Routing via multiple data distribution centers.

5.2 Supporting Peer-to-Peer Interactions

We start this part of the discussion by considering an analogy. netflix.com is a company that allows their customers to rent movie DVDs that are delivered by the postal system. 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 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 the company's operating costs. This analogy is illustrative for our distance learning project, but the issues here are more complex.

5.2.1 Centralized Routing

Let us consider the examples illustrated in Figure 4. Figure (a) is the centralized alternative. In this figure, the students at site A submit homework, some of which is forwarded to teacher X for grading, and some of which is forwarded to teacher Y. (The submission could span two different subject areas.) Teacher Y returns some graded homework to students at site B. And teacher Z returns graded homework to students distributed at sites B and C. Note that the communication events (denoted by the arrows) could be utilizing any of the communication channels discussed in earlier sections, including the Internet or the postal system.

While the disadvantage of this centralized routing approach is obvious, this scheme actually has an important advantage: for example, even though teacher Z needs to return material to multiple student sites, she only sends a single “package” of data to the central server, which acts as a “switch” that demultiplexes data from incoming packages and re-multiplexes data onto outgoing packages. For communication effected by mobile storage devices delivered by the postal system, the central server performs “gather-and-scatter” operations—it copies data from incoming storage devices to outgoing devices, so that each site only receives one storage device and sends one for each mailman visit. For example, when the graded homework destined for students at site B arrives in separate storage devices from teachers Y and Z, the central server copies the data from these separate devices onto a single outgoing device destined for site B. In effect, data routing occurs both digitally and “on foot:” digitally when data is copied from one storage device to another at the central server, and on foot when a storage device is carried to and from the server by a mailman. The server also enforces security policies so, for example, only data that originates from authorized senders and is untampered with is forwarded.

5.2.2 Peer-to-Peer Routing

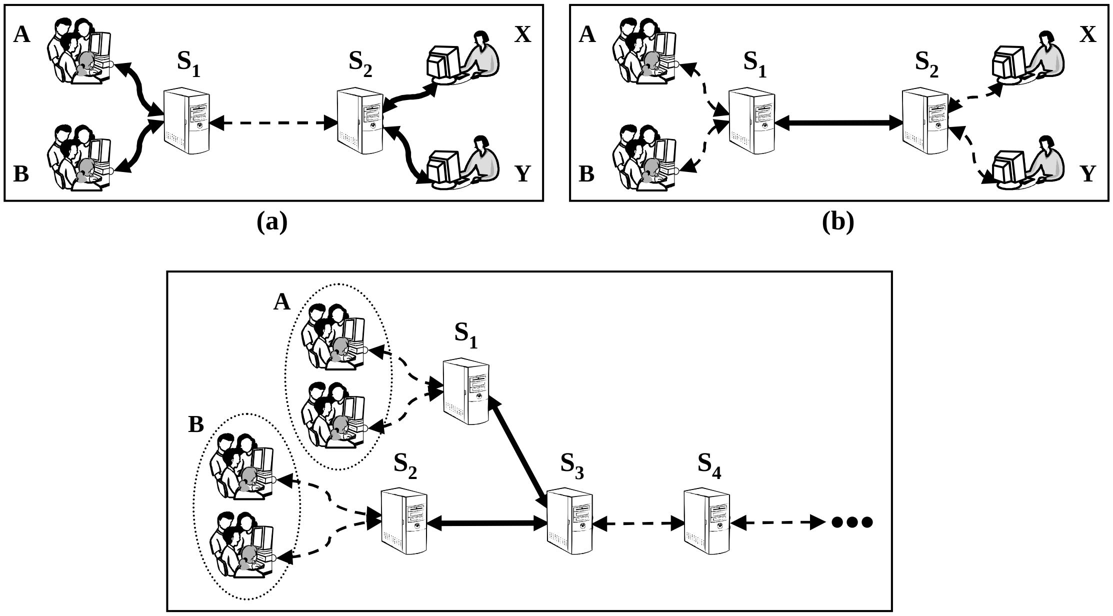
Figure 4(b) illustrates an “opposite” approach. The role of the central server is limited to the coordination of routing decisions: it does not participate in data forwarding. In this figure, the student site A consults the central server to determine which teachers should grade A’s submission. Then site A prepares two “packages,” one for teacher X and one for teacher Y. If netflix had used this strategy, the analogy would be a usage scenario which requires a netflix customer who desires to return his DVD to consult the netflix web site to print out a shipping label; this shipping label identifies the next customer on the waiting list who is interested in receiving this DVD. (netflix would also need to account for potential mistakes or misbehavior by customers had it adopted this strategy.)

Data routing is potentially much more efficient than that in Figure (a). On the other hand, when mobile storage devices are used for communication, and if we consider the use of these mobile devices in a general large-scale publish/subscribe system, a disadvantage of this approach is that a site could receive or send a large number of storage devices per mailman visit, which could become an administrative and cost burden. (In some sense, this is analogous to the inefficiency involved in the use of multiple unicast messages to implement multicast.) For a modest-sized system, however, this approach can be an attractive approach as it demands the least from a shared infrastructure.

5.2.3 Multiple Distribution Centers

Figure 4(c) illustrates a compromise alternative. Instead of employing a single data distribution center of Figure (a), we employ multiple netflix-like distribution centers that are geographically distributed. As is the case in Figure (a), when mobile storage devices are involved, each site only needs to send one storage device toward the closest data distribution center per mailman visit. Data is recopied among devices at the distribution centers. Each site may receive multiple devices per mailman visit, as many as the number of distribution centers. (Or alternatively, a site may send multiple outgoing packages but receive only one incoming package per mailman visit. Or alternatively, a site may employ a mixture of these approaches and send and receive multiple packages per mailman visit. In all cases though, the number of packages involved per mailman visit is limited by the number of distribution centers.)

As is the case in Figure (b), the geographically distributed distribution centers allow some degree of geographical awareness in routing decisions. The distribution centers communicate among



(a) (b)

(c)

Figure 5: Examples of sequentially composing a communication path with hops of different technologies. The solid arrows denote communication effected by mobile storage devices carried by the postal system; the dashed arrows denote communication carried by the Internet.

themselves to coordinate routing decisions. The latency achieved under this alternative, however, 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, while less urgent packages are routed through the data distribution centers to minimize the number of packages.

5.2.4 Transporting Data Across Multiple Distribution Centers

The routing examples shown in Figures 4 (a)-(c) are inspired by how an online DVD rental business might work today and they share the following characteristics. (1) Data transport takes one or two (overlay) hops: two hops when data is routed through a data distribution center (a and c) and one hop when data is routed directly from the source to the destination (b). (2) There is no data exchange among data distribution centers: only Figure (c) employs multiple data distribution centers and the communication among them is restricted to routing information exchange. (3) We have not specifically stated what communication channel through which data is transported: the solid arrows in the figure may denote the Internet, mobile storage devices carried by the postal system, or *parallel* use of both mechanisms (as described in Section 5.1.3).

In this subsection, we loosen the above restrictions and examine options of transporting data across multiple data distribution centers. In particular, we examine the *sequential* use of the Internet and the postal system as data is sent through multiple hops. The motivations for doing so should become clearer as we examine the examples shown in Figure 5.

In Figure 5(a), S_1 and S_2 are two provincial-level data distribution centers (or "servers") that are connected by an Internet connection with substantial bandwidth. The use of this Internet connection allows certain types of data to be transmitted more rapidly across provinces. A and B are two remote villages in the province served by S_1. Due to the lack of high-bandwidth connectivity into these two villages, they principally communicate with the rest of the world using mobile storage devices carried by the postal system. In this case, the postal system helps solve a problem that is similar to one commonly referred to as the "last-mile problem"—the difficulty of extending connectivity to the fringe households.

In Figure 5(b), the roles of the Internet connection and the postal system are, in some sense, reversed. Here, S_1 is a village-level server. Due to its remoteness, S_1 communicates with the rest of the world via mobile storage devices transported by the postal system. An inexpensive local area network (that is possibly wireless) may cover several local sites within the village. This arrangement allows a certain amount of communication and storage infrastructure (represented by S_1), including the postal system-based communication effort (for handling mobile storage devices by mailmen), to be shared across multiple local sites (represented by A and B).

In Figure 5(c), we see a mixture of the approaches employed in Figures (a) and (b). S_1 and S_2 are two village-level servers, each serving their respective local area networks (A and B). These two remote village servers communicate with a provincial-level server S_3 using mobile storage devices delivered by the postal system. S_3, a well-connected server, in turn, may communicate with other servers (such as S_4) via the Internet. The technique of mixing the use of the Internet and the postal system channels that we have described here is orthogonal to that discussed in previous sections. In Section 5.1.3, we exploit the *parallel* use of these channels: multiple channels exist between a pair of communicating parties, and we choose the "right" one to send request/reply messages or send different versions of the same content down multiple channels simultaneously. In contrast, in this section, we discuss the *sequential* use of these channels: a single communication path may be the concatenation of multiple hops, each of which may be a channel of a different technology.

5.2.5 Application-Specific Routing Decisions for Asynchronous Communication

In this section, we consider how data (such as homework assignments that are to be graded) is routed among participants (such as students and graders). What makes the routing problem interesting in our context are three interacting issues: (1) the need of optimizing for distance learning-related application-specific routing metrics; (2) the potentially long latency of asynchronous communication channels; and (3) the abundance of storage and bandwidth of some of these asynchronous communication channels. Before we examine some of the optimization techniques, it is important to note that automation does not necessarily conflict with "people-centric" policies. For example, some students may have preferences on which graders they wish to have; and we may wish to have a relatively stable student-grader relationship. To the extent possible, the system should take such policies into consideration when attempting to find the best way of routing homeworks to the graders.

In traditional routing, one may be interested in optimizing for metrics such as latency. On the other hand, in the context of routing homework submissions to teachers, for example, merely delivering the submissions quickly to a teacher is not necessarily our ultimate goal: instead, a useful optimization metric might be minimizing the time elapsed between the moment when a homework assignment is submitted and the moment when the graded homework is returned to the students. In the examples shown in Figure 4, where multiple equally qualified teachers can all be candidates for receiving homework submissions (potentially via the postal system), the routing mechanism needs to account for factors such as predicted amounts of time needed by the teachers to grade homeworks, their current load (or "queue length"), and the postal system delay incurred to reach these teachers. One challenge is to build the system in a sufficiently generic manner so that all types of application-specific routing metrics can be easily expressed.

Missteps in routing decisions for these long-latency communication networks can be especially detrimental. For example, a teacher chosen by the system to grade a homework assignment falls ill after receiving her assignment in the postal system. By the time the system becomes aware of this problem, it may take days for the postal system to deliver another copy to an alternate teacher. 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. Such replicas can be had virtually for free as long as there is time available for making these copies, due to the abundance of the storage capacity. The first teacher who chooses to grade the homework would send notification messages via a low-latency network (i.e., the 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.)

If there is not sufficient time for making enough replicas to address the imperfections of routing decisions, retransmissions to alternate destinations may need to take place. In the routing architectures shown in Figures 4(a) and (c), data can be "buffered" at the data distribution centers for retransmissions, so that the initial senders do not have to be involved. Teachers can also directly forward their assignments to their colleagues. Again, one of our goals is to build a sufficiently generic and powerful publish/subscribe routing system based on asynchronous communication channels, so that we can account for not only these specific distance learning interactions and scenarios, but also other types of applications.

5.3 Intelligent Human Interface Devices

A new input device that we are currently developing works in the following way. Users write or draw on a white board or a blackboard in the usual way. An inexpensive still digital camera captures a sequence of the board content evolution. An image analysis program analyzes the raw images and identifies the strokes that the users make on the board. The stroke data, instead of the raw images, is the representation that we work with. The data can be more easily transmitted, edited, analyzed, and searched.

This approach has several advantages. It is inexpensive in its equipment needs. It is easy and unintrusive to deploy. The regular blackboards allow a shared group experience, which is impossible on a small screen. The stroke data is more compact to transmit over the network if we choose to. It conveys dynamic evolution without consuming excessive bandwidth. We can adaptively transcode the captured stroke data depending on how much bandwidth is available for transmission. This transcoding can occur in a way that is more intelligent than what is possible for simple raw images or raw movies. For example, we may be able to perform better lossy compression of strokes and stroke evolution in a way that exploits the nature of the stroke data. We could devote different fractions of bandwidth to different regions of strokes.

Because the stroke data can be more easily manipulated, it is easier for teachers to annotate or correct what students have produced. This feedback can happen nearly in real time if we have enough bandwidth to transmit back to the teacher the drawing and the annotated versions (efficiently encoded as differences from the original data, for example). If there is not sufficient bandwidth or if a network connection is not available at all, the students' stroke data can be saved on a mobile storage media, which is to be transported back to the teacher by the postal system for feedback. The teacher can add her annotations and the resulting data can be transported back to the students on a returning storage device.

Stroke data is also potentially easier for an automatic recognition software to analyze. As we have mentioned in Section 3.3.3, limited local intelligence may provide some feedback to the pupils without involving remote interaction with a teacher. In cases where the drawings by students (including both correct ones and commonly made mistakes) tend to fall into a small set of possibilities and are confined to known grid locations, automatic recognition softwares may have a greater chance of success than in more unpredictable contexts. The fact that the stroke data may contain an evolutionary history of the drawings gives more information for recognition softwares to work with. In our colleagues' current work [7], for example, the order of the strokes and the directions of the strokes constitute parts of the feature vector space examined in the recognition process. These patterns may not only allow us to tell right from wrong, but may also shed light on the nature and cause of mistakes, so that the automatic feedback mechanism may provide pre-canned responses. Of course, as we have explained earlier in Section 3.3.3, we only expect these automatic feedback mechanisms to work for a subset of the cases, and human feedback from either the local staff or a remote teacher is needed to handle the rest.

5.4 Related Work

This section is not meant to be an exhaustive survey of all related work: we only selectively discuss some of the most relevant work.

5.4.1 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* and *interactive* data communication channel.

Recent efforts on "Delay-Tolerant Networks" (DTNs) [6, 8, 15] 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 DTNs and our proposed effort. While "postal classes of service" have been mentioned [6], 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 traditional networking connectivity. 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, 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 in Section 5.2.4), these strategies, to a large extent, are merely optimizations, not necessities.

• *Richer resources.* The 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 in Section 5.1, one of the primary technology motivations behind this approach is the observation regarding its huge bandwidth potential. Mobile 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 are "dumb" and "dormant" during transit in the postal system. When they reach their destinations, they are "plugged in," most likely with high-bandwidth wired alternatives (such as USB2 or Firewire). Once such "contacts" are established, they may remain connected for extended periods of time. None of the typical DTN problems in bandwidth, storage space, or power consumption is necessarily a major concern in this 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 focus. On the other hand, some of the research issues that we have discussed in previous sections are unique to our environment. For example, the relative abundance of storage capacity and bandwidth potential of the devices carried by the postal system 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.

For example, as we place new outgoing data on multiple mobile storage devices that are to be carried to a certain destination by the postal system on successive days, we may liberally replicate old outgoing data from previous days on the new outgoing devices. 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.

Another example of a unique aspect of the exploitation of the postal 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 a programming model based on asynchronous "Active Messages," an asynchronous distributed storage system for multiple communication channels, a peer-to-peer interaction system built on top of the asynchronous communication mechanisms, and applications centered around supporting distance learning, are also different from and complementary to the existing DTN research agendas.

5.4.2 The “mimio” Board

A technology that is most similar to our effort of intelligently capturing blackboard drawings is the “mimio” board [13]. It uses ultrasound to track the positions of specially designed markers. Among its potential disadvantages, it is relatively expensive and less natural to the users.

6 Summary: Frequently Raised Concerns and Questions

• *Didn't Bill Gates say that poor people need medicine and not computers [10]?*
Poor people don't need computers but they do need basic education. There is a big difference between simply giving people computers and developing specific and well-defined applications that can change people's lives. We share the belief that simply dumping a pile of equipment on people, in itself, is unlikely to solve any real world problems. In contrast, our goal is improving basic education, a goal that requires the talent and hard work of many people, and the use of many tools, one of which happens to be computers.

• *Shouldn't we devote resources to more pressing problems in the third world than the lack of basic education?*
Peter Bell, the president of CARE, cites improving access to basic education as one of the top three priorities if we were to end extreme poverty [1]. (See Section 2.1, page 2.) (The other two are access to clean water and fighting AIDS.)

• *Do the poor really want to be educated?*
While people in disadvantaged areas may not be well-informed on many issues, an overwhelming sentiment among these people is that education for their children is their only hope for upward mobility. In the recent trend toward privatization of schools in India, for example, a flood of families throughout the country, including those with little cash to spare, struggle to seek out these schools, believing that these for-profit private schools are better than the free public schools[19].

• *Isn't this project about deploying a bunch of existing or dumbed-down technologies? Isn't it the case that there is little challenging computer science research involved?*
No. The resource constraints (such as cost and bandwidth limitations) and the scalability needs of the envisioned system pose significant challenges and would require significant innovations. Section 3 gives an overview of some of the technical innovations and Section 5 spells out some of the details. And there is no doubt that many more research issues (beyond those described in this document) would emerge as we gain more experience and as more researchers join the effort.

• *Can computer scientists do this alone? Can Princeton do this alone?*
No. We believe that technology is only a part of the story. We plan to collaborate with a variety of organizations, both within and outside of Princeton. Section 4 describes some of the potential collaborators. We plan to talk to many more.

- *How do we deal with the poor state of infrastructure in certain disadvantaged areas, such as the lack of electricity?*
Parts of our proposed technical innovations are specifically designed to circumvent the need for a well-developed networking infrastructure. See Section 3.3.1 and Section 5.1. We expect to set up local school sites at places such as local village centers or town centers, where resources such as electricity are available. We do not expect electricity to be universally available in each household, and at least initially, we do not expect to set up local student sites in individual households. If electricity is not available even at locations such as village centers, we will study alternatives such as making available generators or inexpensive solar-powered solutions [11].

- *Can computers ever be better teachers than humans?*
See Section 2.3 (page 3). 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.

- *What is the ultimate limitation on scalability?*
We plan to carefully control the cost of equipment that is deployed at each local school site. We do not expect the equipment cost to be the ultimate bottleneck. We do not expect the distributed software system scalability to be the bottleneck either if we correctly design and build the system. (See Sections 3.3.2 and 5.2.) We conjecture that the number of teaching staff who provide various forms of interaction with children may be the ultimate scalability bottleneck. These may include both volunteers who donate their efforts for free and professionals who teach through the distance learning mechanism for a living. As we have mentioned earlier, the goal of the system is to amplify the utility of this precious human resource.

- *If human resources are the limiting factor in both a traditional system and the proposed distance learning system, how is the latter able to reach a larger population of children?*
We expect to reach a much larger population of children with the proposed distance learning system than without. Our optimism stems from the following factors. (1) *Higher productivity.* See Section 2.3 (page 3). For example, for the teachers who are willing to travel to remote regions to serve the needy children today, distance learning allows them to spend their time more productively by not wasting time on fruitless activities such as traveling. (2) *More attractive working conditions.* For teachers who are enthusiastic about a career dedicated to helping needy children, but are reluctant to endure the hardship of living in less developed regions, distance teaching offers an attractive alternative. (3) *More flexible time commitment by teachers.* For example, a volunteer may donate one hour per day of her spare time to grade homework through the distance learning system. This is a much lighter commitment than what a volunteer would have to commit to today in a traditional system; so we may be able to attract more participants. (4) *More flexible time commitment by students.* Today, children in some rural areas have to help parents with farm work, which, especially during certain times of the year, may interfere with going to school. Distance learning may allow these children to learn at a more flexible pace and schedule. (5) *Increased efficiency through specialization.* See Section 3.3.2 (page 7). 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. Well-trained and experienced teachers who are good at interacting with students should not be burdened with mundane tasks if we were 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. (6) *Leveraging local human resources.* See Section 3.3.2 (page 7). A local staff and upperclassmen may be recruited to shoulder various responsibilities. (7) *Training teaching staff.* We hope to use the same distance learning mechanism to recruit and train an ever-expanding staff. Of course, we expect a quality-control mechanism to be put in place. (8) *Leveraging other teacher resources.* See Section 4. For example, we hope to attract participation by students of teachers colleges. (9) *Automated feedback.* See Section 5.3 (page 23). We hope that automatic pattern recognition software may be able to provide pre-canned responses in relatively simple situations.

• *Why don't we just do something simple, such as digitizing textbooks, and distributing them on CDs?*
See Section 3.1. Simple approaches such as the one mentioned in this question do not provide much customized and timely interaction. While such a self-guided approach may prove fruitful for mature learners who are self-motivated, self-disciplined, and savvy enough to explore and make the most out of a limited learning mechanism, it may not work as well for young learners, who may need closer supervision and more personal and immediate interaction with teachers.

• *Do we have to address a copy protection issue when disseminating our teaching material?*
We expect to follow an "open source" model for the teaching materials that we develop and use [3].

• *How are storage devices carried by the postal system different from existing "Delay-Tolerant Network" efforts such as those involving WiFi-enabled buses?*
See Section 5.4.1. While "postal classes of service" have been mentioned [6], 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 traditional networking connectivity in a general way. The postal system is a truly global and mature "network" whose characteristics are very different from typically localized and resource-starved "challenged networks." As a result, the research issues that we face are also different.

• *Does this project have any relevance in the "first world?"*
Yes. First, the lack of teaching resources is not a problem that is unique to developing countries—we expect a distance learning mechanism that amplifies the reach of quality teachers to be applicable to developed regions as well. Second, the computer science technologies (Sections 3.3 and 5) that we plan to develop should be applicable to a wider array of applications than distance learning. Third, we expect that graduate students who base their research work on this project to hone their skills in a valuable and unique real world setting.

References

[1] P. D. Bell. Where the end of poverty begins. http://www.princeton.edu/WebMedia/lectures/, February 2003.

[2] The world factbook. http://www.cia.gov/cia/publications/factbook/, 2003.

[3] Creative Commons. http://creativecommons.org.

[4] D. E. Culler, A. Dusseau, S. C. Goldstein, A. Krishnamurthy, S. Lumetta, T. von Eicken, and K. Yelick. Parallel Programming in Split-C. In *Proc. of Supercomputing '93*, November 1993.

[5] X. Ding. Rural area distance learning: contributions today and benefits for all future generations. http://-www.chinaonlineedu.com/media/200314/zl3.asp, 2003.

[6] K. Fall. A delay tolerant networking architecture for challenged internets. In *Proc. of ACM SIGCOMM 2003*, August 2003.

[7] J. Ge and R. Schapire. Online chinese handwriting recognition. http://rywang.cs.princeton.edu:8000/Projects/-03sIND/submit/jge.doc, May 2003.

[8] A. A. Hasson, R. Fletcher, and A. Pentland. DakNet: A Road To Universal Broadband Connectivity. http://-courses.media.mit.edu/2003fall/de/DakNet-Case.pdf, 2003.

[9] P. Juang, H. Oki, Y. Wang, M. Martonosi, L.-S. Peh, and D. Rubenstein. 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] K. Keniston. Grassroots ICT Projects in India: Some Preliminary Hypotheses. *ASCI Journal of Management*, 31(1&2), 2002.

[11] P. Kennedy. Necessity Is the Mother of Invention. In *The New York Times Magazine*, November 30, 2003.

[12] T. Li. Bringing about rapid improvement to elementary and middle school education in midwest rural areas in China. http://www.edu.cn/20030424/3083447.shtml, 2003.

[13] mimio Virtual Ink. http://mimio.com.

[14] J. Pomfret and P. P. Pan. Chinese Premier Presses U.S. on Taiwan, Trade. In *The Washington Post*, page A01, November 23, 2003.

[15] R. Shah, S. Roy, S. Jain, and W. Brunette. Data mules: Modeling a three-tier architecture for sparse sensor networks. In *IEEE SNPA Workshop*, May 2003.

[16] S. Sobti, N. Garg, C. Zhang, X. Yu, A. Krishnamurthy, and R. Y. Wang. PersonalRAID: Mobile Storage for Distributed and Disconnected Computers. In *Proc. First Conference on File and Storage Technologies*, January 2002.

[17] S. Sobti, N. Garg, F. Zheng, J. Lai, Y. Shao, C. Zhang, E. Ziskind, A. Krishnamurthy, and R. Wang. Segank: A Distributed Mobile Storage System. In *Proc. Third Conference on File and Storage Technologies*, March 2004.

[18] T. von Eicken, D. Culler, S. Goldstein, and K. E. Schauser. 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)*, pages 256–266, May 1992.

[19] A. Waldman. India's Poor Bet Precious Sums on Private Schools. In *The New York Times*, page 1 (Section A), November 15, 2003.

[20] R. Wang, N. Garg, Y. Shao, E. Ziskind, S. Sobti, F. Zheng, J. Lai, and A. Krishnamurthy. A Peer-to-Peer Mobile Storage System. In *Business Briefing: Data Management and Storage Technology 2002*. World Markets Research Centre, October 2002.

[21] Y. Wei. Improving education in western provinces using information technology. http://online-edu.org/article/-article/3154.html, 2003.