

A SCALABLE ENABLING IT INFRASTRUCTURE FOR DEVELOPING REGIONS (ICT4B)

There are thousands of stand-alone projects that aim to bring information and communication technology (ICT) to developing regions, but nearly all of them depend on existing hardware and infrastructure developed for affluent regions. As such, these imported technologies fail to address key challenges in cost, deployment, power, and support for populations with low levels of literacy and multiple languages. This project, *ICT4B* (Information and Communication Technology for Billions), develops the key technologies and infrastructure to enhance these projects, and enable new, previously intractable projects.

The proposed work directly addresses these challenges with novel technology, and tests the impact through the use of two real-world pilot deployments in two different developing regions. The purpose of these two deployments is not just to understand and address two specific ICT applications, but also to demonstrate and validate that the underlying ICT architecture and associated technologies are truly enabling for developing regions. Towards this end, the multi-disciplinary proposal includes two faculty and three students from the social sciences to ensure that the work enables real-world solutions. In addition, supporting partners, such as Intel, HP, UNDP, IIT Delhi, Markle, and Grameen Bank will provide project guidance and on-site help with deployments.

Intellectual Merit: The technology strategy is to attack directly the key challenges of cost, power, deployment, support and literacy. ICT4B will provide order-of-magnitude improvements in device cost, infrastructure and networking cost, and power consumption. Key deliverables include 1) novel low-cost, low-power devices, 2) a new approach to low-cost networking based on intermittent connectivity (rather than persistent connectivity as assumed by the Internet), 3) a user-interface toolkit that supports poor literacy through novel speech recognition, and supports a variety of sensors for environment and health applications, and 4) a three-tier architecture that uses proxies and data centers to support low-cost devices with more functionality, easier development, over-the-wire reprogramming, and usage monitoring for social science research. The expected 10-100 times reduction in device cost stems from the co-design of the devices and infrastructure, system-on-a-chip integration, and the extensive use of open standards. The 10-100 times reduction in device power usage stems from the use of low-power circuits, the need for less CPU power due to help from the infrastructure, and novel low-power intermittent networking. The 10-100 times reduction in infrastructure cost comes largely from 1) the focus on intermittent networking, which avoids the substantial costs due to real-time constraints and persistent connections (both present in cellular telephony systems), 2) extensive sharing (across applications and possibly even regions), and 3) novel architectural approaches to user and system support in the field. From a social science perspective, this project will create unique opportunities to develop and test propositions about the role of digital technologies in helping to overcome long-standing and intransigent development problems.

Broader Impacts: ICT4B initially targets five application domains for the ICT framework: commerce, health, education, government and location-based services, as presented in detail below. Although the two pilot deployments will have substantial impact for the (developing region) users of those applications, the true impact of the proposed work is its ability to enhance nearly all of the existing developing-region IT projects, and to enable new projects that would otherwise be intractable. The nature of this leverage may vary: some groups will directly use the artifacts such as the low-cost devices, some groups will exploit the architecture and networking protocols with their own devices, while some groups will leverage the core technologies for low-cost displays, low-power radios, or user interfaces and speech recognition. This project is unique in that makes nearly all of the existing IT projects better: for example, reducing device costs and improving network usage both imply that many more users can achieve the benefits of any given project. ICT4B does not *require* infrastructure for every rural village in the developing world from the outset (over 500,000 villages in India alone). Rather, the architecture will be economically viable, and deployable as such, often in the micro-franchise model of Grameen Bank. Even with only a 5% penetration of a large and developing population, the effect would be very significant. Finally, ICT4B will inspire students, universities and corporations to take bolder, more direct, and better-integrated approaches toward IT for developing regions.

1. PROJECT DESCRIPTION

Information and Communication Technology (ICT) has had profound effects in developed countries and for elites everywhere. But it has arguably increased the opportunity gap between privileged groups and the four billion people who have no access [Cas01]. Although progress has been made in improving the human condition in some developing countries, enormous challenges remain. In 1999, 2.8 billion people lived on less than \$2 a day, with 1.2 billion subsisting on less than \$1 a day.

We propose to initiate a program, *ICT4B* (Information and Communication Technology for Billions), aimed at making the benefits of ICT available and useful for all areas of the world. Our interdisciplinary team will focus on developing a hardware/software infrastructure explicitly designed for the physical, political and economic realities of developing areas.

Information technology is not a silver bullet for long-standing development issues. Some argue that developing countries need investments in healthcare, the provision of clean water and low-cost energy ahead of an effective ICT infrastructure. We believe that is a false choice, and misses the extent to which information technology can help meet basic human needs. The fact is, all such investments play a important role, but a viable and effective ICT infrastructure provides a necessary substrate for economic development, more effective local government, education and training, the effective deployment and support of other technologies and know-how (like those listed earlier), as well as the promotion of stable economies and democratic values [GTR02].

The transformative potential of ICT creates new kinds of opportunities to gain leverage on development problems that have been intransigent. The potential lies within the tools that ICT provides to *all* economic and political sectors. These tools can affect every activity in which organization, information processing, or communication are important. They change what can be achieved and how it can be achieved across a wide range of industries and public enterprises. Most important, they drive changes in ideas about ownership, property, and control – the way in which governing institutions order economies in the broadest sense of that term. Using ICT to challenge suboptimal institutions and particularly corrupt and inefficient “markets” provides a rare chance to fashion new, socially beneficial equilibria that contribute to the broader development goals of countries and peoples.

This argument is implicit in the main conclusions and recommendations of the G-8 Digital Opportunity Initiative Report [DOI01] and similar reports from the World Bank and UNDP [WB01, UN01], as well as the 2001-2002 Global Information Technology Report [GTR02]. ICT is of course not a substitute for poverty alleviation and governance reform, but instead a broader framework for incorporating these goals into a medium-term strategy for enabling development.

2. OVERVIEW

To date there are thousands of projects that aim to use ICT to help address some specific application in some specific region. The World Resources Institute alone lists over 600 projects in developing regions that are IT related [WRI03]. There are many organizations that selectively collect data on IT around the world, particularly with regard to developing countries. Many of these organizations, such as the Benton Foundation, Global Knowledge, Development Gateway, and Bridges.org, maintain large online databases of case examples. The emphasis of these organizations is on projects related to Internet software and access, rather than innovative hardware, software or infrastructure solutions. A recent search on these databases uncovered only four examples of new hardware among thousands of cases.

The fundamental distinction of this proposal is that we aim to enhance all of those existing applications with innovative hardware, software and infrastructure, enable new solutions that would otherwise be intractable, and inspire students, corporations and governments to attack these critical problems in a new way. With the help of our partners, we plan to develop and deploy (in two distinct regions) a novel and scaleable ICT substrate, with an additional goal of ensuring an economically viable long-term deployment solution. We will enhance and enable these applications in three ways:

1) Novel Core Technology: ICT designed for well-developed economies is unlikely to match the power, cost and communication needs of developing regions: there is little short-term financial incentive to do so. Thus we will develop new technologies in several critical areas that we believe are unlikely to evolve on their own:

- **Low-power intermittent networking:** The current focus for commercial networking infrastructure is real-time continuous communication, typified by the wired and cellular telephone infrastructure and satellite Internet access. Although useful, we believe that many information-related applications and very effective asynchronous communication could be designed around intermittent-connectivity networking (ICN) instead. Unlike ICN, continuous, real-time communication is extremely expensive to provide and requires substantial power. ICN messages avoid real-time constraints, are easy to multiplex (saving cost), and can be delayed or batched as needed to reduce power and cost. We will develop both novel protocols and low-power circuits for this class of communication including reliable or prioritized delivery and return-receipt.

- **Flexible Low-Cost Devices:** The current set of “low-cost” devices used for these applications are essentially low-cost PCs that typically cost more than \$100 each; the Simputer is a good example [Iye01]. In contrast, we are targeting a family of devices that cost as little as \$1 each. To achieve this 100x difference we will: a) make the devices network centric, so that we can offload some processing and storage to shared infrastructure, b) develop a new system-on-chip architecture that integrates computing, storage, and wireless networking and thus greatly reduces the component count and cost, c) reduce the power consumption by an order of magnitude, which reduces battery/power costs, and d) develop low-cost low-power displays based on PDLCs and Inkjet printing technology. We will exploit our work with wireless sensor networks for these devices [JSW+00, GLvB+03, RAK+02].

- **User-Interface Toolkit:** There has been relatively little work in the development of effective user interfaces for these applications, especially in an environment with low literacy and with a wide range of languages and dialects. We will develop a toolkit that enables these applications by providing support for small-vocabulary dialect-specific speech recognition and generation, which is critical for regions with low literacy and obscure languages or dialects. We envisage a multi-media communications interface, where images and gestures, as well as simple sketches, can be used to enhance and complement the speech interface even further. The toolkit will also handle interaction with a variety of additional sensors that may be useful for a range of applications, including health and environmental sensors (e.g. water testing), and ultra low-cost CMOS digital cameras [Pix03].

2) Shared Infrastructure: Because all of the current applications are built independently, there is essentially no shared infrastructure and almost no shared learning. For any given application, the cost of infrastructure is hard to justify; however the existence of shared infrastructure would benefit nearly all of the applications. For some applications the benefit is simply lower cost or more users, while other applications are simply not tractable without shared infrastructure. We will co-design the devices and infrastructure together to reduce the overall cost of the system: in general, moving functionality into the infrastructure reduces its cost by 10-40 times, primarily due to amortization and sharing of costs [Bre01]. In addition, we will deploy and operate the shared infrastructure with our partners, with the hope and expectation that other groups will be able to use it for their own applications. Finally, with a shared and decentralized infrastructure, specially trained local users can provide local and therefore low-cost user and system support. Of course, the system must be designed with such a support model in mind and suitable basic training provided (via the same system). The experience of the Grameen “phone lady” [Coh01] and the success and growth of distributed call centers in developing countries (especially India) suggest that this is likely to be a successful approach.

3) Faster Research and Development: Our use of shared infrastructure enables a third fundamental benefit for these applications, which is a much better platform for ongoing research and development. First, researchers can monitor the application in action, using the intermittent networking to understand system usage and behavior. This has historically been difficult because the devices were disconnected. Second, we will develop the ability to reprogram the devices remotely (over the network), which means that we need not collect and redistribute devices to evolve the system. This will greatly increase the rate of evolution and facilitate broad experimentation. In turn, we expect these benefits to lead to better overall solutions, a higher success rate, and broader adoption.

VALIDATION: Although we can easily evaluate these enabling technologies in isolation, it is difficult to validate that the framework actually enables applications and truly empowers users in a developing world. Our strategy for validation is to *build and deploy two novel applications* that use the framework in two different regions of the world. In addition, we expect that other projects (e.g. a number of the projects within CITRIS) will also use the infrastructure within the five-year lifetime of the project.

We are targeting five application domains for the framework: commerce, health, education, government and location-based services; we cover these in more detail below. Although we will design the framework for all five application categories, and we will consider applications in all five areas, we will select two specific applications for deployment. We outline the criteria for selection in Section 2. This is a true multi-disciplinary project: the choice and evaluation of applications will be led by social scientists (Weber and Ray), and we expect these applications to generate at multiple social-science Ph.D. degrees for their design and evaluation.

The first of the two deployments will be in India, starting in year three, with updates in years four and five. We have chosen this location based on our partners, including HP, which is supporting an i-Communities project in India, and IIT Delhi. These partners have substantial local presence and knowledge that make a real deployment tractable, and HP is a core CITRIS industrial partner. The second deployment will start in year four and will be in a region outside of Asia (for diversity) and will most likely be closer to, or even within, the United States. We will decide on the second deployment during year two, and it will also depend on local partners.

SCHEDULE AND MILESTONES

Year 1	<ul style="list-style-type: none"> • Develop one complete prototype with all three layers. Device will be based on commodity technology (e.g. Bluetooth) for the device-to-proxy networking. • Investigate possible social placements for pilot deployment (in India) • Investigate application domains, using classes and conferences, and decide on application for the first deployment (in India). • Demonstrate intermittent networking protocols. • Demonstrate basic low-cost display technology • Define UI toolkit requirements
Year 2	<ul style="list-style-type: none"> • Development of integrated radio/storage/bus for devices • Demonstration of small pixel arrays (display) • Deployment of local testbed for testing • Demonstration of applications using the initial UI toolkit, including speech recognition • Finalize India deployment plans • Decide on short list for deployment 2 (both application and location)
Year 3	<ul style="list-style-type: none"> • First deployment in India, and initial evaluation • Finalize second deployment plans, and second iteration for India deployment • Demonstrate applications using everything except new displays • Demonstrate full UI toolkit and speech support • Release of intermittent networking protocols • Demonstrate complete display functionality including low-power architecture
Year 4	<ul style="list-style-type: none"> • Second deployment and initial evaluation • Version 2, deployment in India, advanced evaluation • Demonstration of single-chip devices • Demonstration of wireless networking with self-sustaining power train (no batteries) • Demonstration of complete low-power displays
Year 5	<ul style="list-style-type: none"> • Fully integrated single-chip device with low-power display driver • Final evaluation of India deployment, both technical and application oriented • Version 2 of second deployment, advanced evaluation • Overall evaluation of framework for enabling applications • Transition of active services and infrastructure to partners

3. APPLICATIONS [WEBER, RAY, BREWER, CANNY, DEY, FELDMAN]

Because of its broad goals, the ICT4B project cannot be focused on a narrow range of applications; at the same time, it makes no sense to design a system without applications. Our approach to bridging this gap is to start by specifying five application domains that are both important and also feasible for the systems and environments envisioned in the project. These are: commerce, health, education, government, and location-based services. In all of these, the availability and exchange of information is a critical factor,

and we expect that many applications in each domain will map well onto our model of shared infrastructure and intermittent networking.

We realize that there are much larger efforts than ours working on each of these areas and we are engaged in learning from them. As one step we have organized a workshop on “Enabling Digital Technologies in Emerging Markets” for Summer 2003, with support from the Ford Foundation and Leverage Foundation. We also plan to participate in a conference organized by CMU on IT for sustainable development. We are also planning a new graduate seminar course at Berkeley, to be first offered in the Fall 2003 semester, to collect and evaluate existing ICT architectures and approaches in the developing world, to determine a detailed list of projects and technology alternatives, and to ask our students to evaluate them in detail. In addition, several of our partners have existing projects and experience to provide guidance, including the programs at HP i-Communities, IIT Delhi, Markle, and Grameen Bank. We plan to make the seminars and course materials available via webcast and for later viewing as well, using the CITRIS multimedia facilities, and will include our international partners.

A unique aspect of our pilot deployment process is that we plan to choose the specific technologies and the application areas to be implemented in concert with their social placements. Socio-economic evaluations and environmental impact evaluations are routinely performed when new projects or new technologies are implemented. But most such evaluations focus on *overall* welfare increases – they rarely focus on reducing existing social and economic inequalities [Git82]. As a result, all too many technical innovations—from high-yielding seeds to agricultural machinery, and from medical advances to educational tools—have exacerbated the social differentiation between men and women, rich and poor, elite- and low-caste [Mei95]. But central to our definition of ICT4B is equity of access to the technologies that we deploy. We therefore propose to conduct *ex ante* evaluations of several sites in the first year of our research program – so that we eventually place the devices within communities that are willing to extend access to the traditionally under-privileged (such as women or minorities). The social scientists on our research team are experienced in these evaluations, through which we can learn, among other things:

1. Which among the applications being developed will meet the felt and articulated needs of the community?
2. What are the baseline characteristics of the community (in terms of income, access to markets and healthcare, literacy levels, etc)?
3. What economic, health or educational benefits are likely to accrue from these technologies, and to whom?
4. How can the technologies be implemented such that women or under-privileged minorities are not excluded?

The co-evolution of social appraisal and technical choice will minimize the risk of failure, maximize the chances for socio-economic development of the truly marginalized, and provide a comprehensive learning experience for the future deployment of decentralized technologies.

Strong partnerships with the voluntary sector will be extremely helpful in the site-selection and pilot deployment processes. Experience has shown that mature civil-society leadership and community capacity are crucial for the success of development interventions. With decentralized technologies and numerous possible end-users, the need for a broad base of knowledge is especially critical [DR02]. One of our researchers (Ray) has worked closely with health, education, and sustainable development NGOs in India for over 15 years, so we are confident of our ability to choose effective partner NGOs and/or implementing agencies for our first deployment.

There appear to be two generally effective modes of organization for broad-based technology initiatives in developing regions. One is the franchise model, such as village phones, led by our collaborator, Grameen Bank. The second is the enhancement of a current service with existing public or private support, such as health or education. We believe both can lead to sustainable deployments.

The following preliminary analysis of the key application areas has shaped our initial design decisions, but we realize that we have a great deal more to learn about all of them.

COMMERCE-ENABLING APPLICATIONS FOR EMERGING MARKETS

A key feature of less-developed economies is that they lack essential parts of a market platform—the institutional infrastructure on which trading would take place in a more developed economy. Market making is a prerequisite to price discovery. Pareto-improving transactions fail to occur in developing economies because demand and supply may be disconnected by distance, by information opacity, by language, by politically imposed barriers to trade, etc.¹ The parties to a potential transaction need to know that each exist, that they have a potential deal of mutual benefit, and that they can reliably exchange whatever is going to be transferred in both directions. Even in very poor developing economies, a transaction often requires some asynchronous exchange of value and may involve third parties as de facto financiers or insurers. Allowing these people to find each other, coordinating the time and place of their physical interaction, and creating a record of their transaction are all basic features of market making. All of these actions can exploit intermittent networking and shared infrastructure. The use of shared infrastructure also enables the establishment of reputations over time, which facilitates relatively anonymous transactions and thus expands markets beyond the very limited local circle of trust.

HEALTH

The health sector, which includes individual care, health education, and public health, is the prototypical example of a domain in which information can be a crucial factor. Many common health problems in developing regions involve dangerous but highly treatable diseases for which recovery is excellent if suitable treatments are delivered in a timely fashion. For the common problems such as diarrhea, cholera, and malaria, human diagnosis is relatively reliable from a few observations. Automated diagnosis systems using probabilistic reasoning are highly evolved and achieve near-human accuracy on a limited range of diseases [BB98]. Such systems, if appropriately integrated with referrals, could provide highly effective treatment recommendations without complexity of a physician appointment. More complex conditions necessitate physician care, but we can facilitate referrals and location of help.

Many of the health problems are environmental. By deploying environmental monitors, equipped with GPS-based self-location and intermittent networking, we believe we may be able to track the spread of specific health-related problems. This will enable health workers to more effectively monitor and predict the spread of disease and better focus scarce preventative resources and techniques where they can be most effective. The PIs have significant experience with wireless sensor networks and environmental monitoring [MPS+02, Can02, RAK+02, DR02].

Further out, we envision extending Berkeley's work on telepresence for live remote consultation [PC98]. This work is currently in use in the NSF-sponsored NEES project on campus [NEES]. This project involves simultaneous live consultation, unstructured discussion, and sensor data gathering and analysis. Similar techniques can be applied to remote primary health care, with the physician in the loop. Simple sensors for basic vital signs are commercially available, although it remains a challenge to develop these at extremely low cost. Because of the need for real-time communication and complex sensors, we view this extension as an application that would use the shared infrastructure, base devices, and UI toolkit, but that would probably be outside the scope of intermittent networking.

LEARNER-CENTERED EDUCATIONAL AND TRAINING ENVIRONMENTS

Education and training, with or without technical mediation, is a great challenge in developing countries. Over the past few years, we have seen significant success with on-line training systems for specific technologies and application areas (e.g. Cisco). Many believe the augmentation of local communities with specific training and new skills will enable huge (relative) economic gains in those regions (e.g. remote call centers, effective access to distribution for rural products and services). *An effective ICT substrate is critical to such remote training and service applications.* As Venture Capitalist Vinod Khosla puts it, "I suspect that by 2010 we will be talking about [remote services] as the fastest growing part of the world

¹ *Pareto improving* means simply that at least one party to the trade is made better off and no one is made worse off by the trade; in other words, there is an aggregate increase of utility without anyone being hurt.

economy, with many trillions of dollars of new markets created. There is no reason why services like bill processing should not be done independently of a company” [Koh00].

It is clear that technology alone cannot have a significant impact on education, but new learning practices especially learner-centered techniques, when combined with appropriate technologies have shown dramatic improvements at the middle school through to university level [SC97, LC03]. The experience of rural schools in developing countries such as India has shown that even children from semi-literate families show a natural enthusiasm and affinity for PC-based learning. Education-oriented NGOs in Andhra Pradesh, Uttar Pradesh and Karnataka have found that these children enjoy the hands-on, experiential approach to learning that computer-based applications provide. They are also convinced that computer-based skills are a more direct route out of poverty than traditional classroom-centered instruction, so they are more motivated to learn.²

The particular promise of these techniques in developing regions is that they place far fewer demands on the qualifications and availability of a nearby instructor. They are also much more responsive to an individual learner’s expertise. The problem then becomes one of developing content to match the students’ language, cultural background, and level of preparation. Berkeley CS faculty in cooperation with the School of Education have been developing rapid-authoring tools to enable less-skilled educators to develop inquiry-based content that is adapted to the needs of their student audience. The tools make use of pattern techniques to allow the instructor to reuse courseware developed, say, in another language, to their own students’ needs. Level of preparation will change too, but the basic structure of inquiry-based courses uses the same elements, and only individual modules need to be adapted to learner preparation. Its success from middle school through college level suggests that it is broadly applicable.

The infrastructure and display technology outlined in this proposal would keep distribution costs extremely low, and intermittent networking and shared infrastructure form a natural approach for this domain (analogous to courses by mail but faster and more interactive).

E-GOVERNMENT

Research supported by the World Bank has concluded that e-government can provide concrete benefits to governments and citizens in developing countries [WB01]. Potentially positive outcomes include: (1) increased access to government information and services, (2) cost savings and service improvements; (3) greater interaction between the public and government officials; (4) increased transparency and reduced opportunities for corruption; and (5) new development opportunities for rural and underserved communities. For example, in India, the Madhya Pradesh State Initiative has created an intranet that reaches remote villages in rural India. For ten cents, farmers can obtain copies of land titles that previously cost \$100 when purchased from corrupt government officials. Villagers can report broken pumps, lost pension checks, or a sick teacher – with the state guaranteeing a reply in a week [CDT02]. Also in India, the Central Vigilance Commission has started publishing on the Internet the names of officials suspected of soliciting bribes, and the disciplinary actions taken against them [CDT02]. This increases accountability and confidence in the system, since citizens can immediately tell whether corrupt officials are being punished. In Namibia, citizens can comment on legislation under consideration before the Parliament and regional governmental organizations [CDT02]. President Vicente Fox of Mexico, as part of his “e-Mexico” program, has already declared that transparency through technology is a key part of his plan for eliminating corruption [Pat02]. All of these examples fit well with our approach; the information is clearly shared, has no real-time requirements, and must reach a partially literate audience.

LOCATION-BASED SERVICES

Much of what we take for granted in the West is a challenge for everyday life in the developing world. Access to adequate potable water each day remains a challenge to many. Rural life depends on other resources like grazing land, and farming is ever vulnerable to the spread of pests and diseases. Natural disasters such as flood and fire still take many lives, in part due to the absence of timely information.

² Case studies can be found on <http://www.ashanet.org> and www.icaonline.org

Sharing information about these critical resources and risks would ameliorate many problems. Location via GPS fortunately uses free infrastructure and is available around the world. “Situation awareness” at particular locations can be generated in two ways: 1) directly from widely deployed sensor data, and 2) through information sharing by communities. The latter in particular uses information derived from already deployed devices, and is economically straightforward. Senior Personnel Canny under existing NSF support is developing novel information-sharing algorithms combining both user and geographic information to answer location-based queries for mobile devices.

In many cases, even basic location information would be novel: there is rarely an available “phone book” of businesses, nor are maps always available or up-to-date. Location-based services can provide timely data about road and bridge conditions, flooding, landslides and other obstacles. Information about the nearest source of good-quality water is always helpful and sometimes saves lives. Information about the location of businesses, seasonal marketplaces, health and other services, can be communally maintained at little or no cost. We believe even basic steps in this area, based on intermittent networking and shared infrastructure, would have tremendous impact, and enable many other applications.

4. TECHNICAL CONTRIBUTIONS

Having looked at the five application domains, we turn to the system design and the technical contributions. There are three key architectural innovations in our approach. The first is the holistic co-design of the devices and the infrastructure, which greatly reduces device cost and thus overall cost, and enables shared information and faster research and development. The second is the focus on intermittent networking as a design principle; it allows a network-centric approach to solutions (rather than device-centric), while allowing significantly lower-cost infrastructure when compared to the existing approach based (largely) on cellular telephony. Third, our focus on a family of communication-centric system-on-a-chip devices enables extremely low-cost and low-power devices.

In addition to architectural innovations, we will deliver novel core technology in the areas of low-cost displays (Section 3.4), user interfaces and speech recognition (Section 3.5), intermittent networking protocols (Section 3.2), and low-power devices with integrated wireless networking (Section 3.3).

ARCHITECTURE: DEVICE-INFRASTRUCTURE CO-DESIGN [BREWER, FALL, RABAEY]

Nearly all approaches to large-scale use of ICT focus on low-cost devices that are either stand-alone or use existing basic connectivity, such as cellular infrastructure. A key innovation in this proposal is to focus first on the shared infrastructure, and co-design the devices with the infrastructure. Of course, this does not preclude the use of existing infrastructure; in fact, it improves the efficiency and capacity of such infrastructure in ways that should extend the life of existing investments. Co-design in turn enables 1) significantly lower overall cost (including device cost and theft), 2) much greater flexibility and reuse, 3) lower energy usage, 4) easier upgrading and less obsolescence, and 5) significantly easier research and development. We cover each of these in turn after a brief overview of the three layers of the architecture.

Figure 1 shows the layers of our three-tier architecture. The top layer, the “device layer”, consists of devices, which may be shared or personal. We expect a very wide range of devices, both in functionality and cost. The middle layer is the “proxy layer”, which is shared at the level of a town or village. Proxies form stable gateways to the larger infrastructure and have better connectivity and power than devices. The bottom layer is the data-center layer, which consists of few widely shared data centers with reliable power and storage, and considerable overall resources. Because these centers can support millions of users using commodity hardware, their cost per user is extremely low. We do not view the use of three tiers as novel, and in fact we have significant experience using this structure for (Western) Internet applications and sensor networks [Bre+02, MPS+02]. However, it has not been applied to developing regions, and the use of co-designed devices (rather than commodity devices) is novel.

This tiered structure significantly reduces overall cost. First, there is substantially more sharing at the lower layers: proxies can support thousands of users, and data centers can support up to a billion users. The extreme sharing leads to very high utilization of components, and more predictable utilization, both of which reduce cost. Typical cost savings from the Internet range from 10 to 40 times [Bre01]. The lower

layers also have much lower operational cost per user, due to shared administration and maintenance. Finally, the fact that devices are less useful without infrastructure support can reduce theft.

In our model, the lower layers support wide variation at the device layer, and reduce the development cost for new applications and devices. Although we believe that each application should be designed top down with a specific device and user interface, we expect all applications can easily reuse the infrastructure layers. In addition, in moving development of features from devices to “devices + infrastructure”, we can exploit better tools and leverage existing code and infrastructure (such as databases), which greatly simplify the development of complex features.

We also expect energy savings due to shared infrastructure: Every task that is moved from the device to the infrastructure not only reduces cost, but also adds battery life to the device. When combined with intermittent networking and low-power design (discussed below), we expect an order of magnitude improvement in device battery life.

A traditional problem with device-centric applications is that they are hard to upgrade: an upgrade requires modifying the device in the field. We have demonstrated over-the-network (software) upgrades to our devices [LC02], and can easily upgrade the bottom layers. This in turn enables more aggressive software (because mistakes are easy to fix), iterative development, and reduced obsolescence.

Finally, because there is a fundamental connection between the devices and the infrastructure, it is easy and natural to monitor how the devices are being used. We believe this it be a significant enabler for researchers that will result in better devices and applications, and thus better adoption. Given this overview, we next examine each of the layers in greater detail.

The device layer will include a wide range of devices, both in cost and functionality. Simple communication devices (like pagers) can be very low cost (< \$1), while devices with complex sensors, such as water or chemical monitoring, will be significantly more. We expect to use the TinyOS operating system and its development language, nesC, for devices, both of which were co-developed by PI Brewer [JSW+00, GLvB+03]. This OS was developed for networked sensors, or *motes*, which have the desired properties of super-low cost (moving towards cents/device), support for wireless networking, and very low power and space usage. For example, the TinyOS kernel is less than 1KB. In addition, we inherit the simulation tools and a wide range of supported sensors. We expect the packaging and display to dominate the cost of the device, hence our focus on low-cost displays and voice-based interaction.

Layer	Hardware	Networking	Number	Users	Cost	\$/User
Device	Small devices	Short-range wireless	> 1 million	1-10	\$1-\$10	\$1-\$10
Proxy	PC Basestation	Cellular, Satellite, Wired	100-1000	100-1000	\$200	< \$1
Data Center	Rack-mount PCs	Full Wired Internet	< 10	>1,000,000	\$100,000	< \$0.10

- Devices range from simple communication devices (like pagers) at the low-end to complex devices with sensors for trained users (water testing). We may also support very expensive devices for certain cases such as remote medical diagnosis.
- Persistent storage is primarily at the data center, but there will be some on the device (e.g. flash memory).
- Proxies have only transient storage for in-transit messages, caching, and broadcast data.
- The network between the proxies and the data centers will vary, and will essentially be whatever is most convenient and cost effective, such as satellite or fixed wireless based on 802.11.

Figure 1: The Three Architectural Layers

The proxy implementation will be based on Linux PCs, possibly using Intel’s X-Scale processor for lower power [Int03]; our partner Intel has already developed a low-cost Linux computer based on this processor (formerly the StrongARM). Another alternative would be to evaluate the applicability of a task-specific programmable processor like the Tensilica X-tensa architecture and it associated development environment; we have been using this approach in some of the research underway in the Berkeley Wireless Research Center with significant success.

Proxies have only transient state, and are best viewed as a combination of a cache for shared information such as market prices, with a gateway that acts as a waystation for network communication with intermittent connectivity. Proxies are also responsible for local computation that assists devices, such as triangulation, image processing, and voice processing. By placing these services on the proxy we reduce device cost and share computing resources for rarely used but expensive operations. We previously developed the first use of proxies to enable devices and have extensive experience in this area [Bre+02], including work that led to US consumer use of proxies for PDAs and cell phones. The partitioning of functionality among the layers is very flexible and will be application specific; in general it is an open research question that we hope to answer. For example, speech processing on the device increases device cost, but offloading it reduces response time and power consumption as well as cost. Adaptive partitions are also possible.

The data-center layer will be based on cluster computing, using clusters of commodity Linux PCs. Starting with the NOW project in 1994, Berkeley has a long track record of using clusters to support live 24x7 services [Bre01, FGC+97]. We expect to develop and operate a single data center at Berkeley that will support all of the deployments abroad. In the future, we would transition these centers to long-term locations that are more convenient for the service operator.

Key deliverables:

- Three-tier architecture with flexible partitioning of application functionality.
- Demonstration of at least 10x cost and power reduction due to offloaded functionality.
- Reliable over-the-network upgrades of devices and proxies.
- Demonstration of enabled social-science research through non-intrusive anonymous usage monitoring.
- Scalability for devices and persistent state (in data center) up to 1B users (via simulated devices).
- Open source versions of device, proxy and data center application code.

INTERMITTENT-CONNECTIVITY NETWORKING [FALL, BREWER]

Recent efforts at providing wide-area data connectivity have focused primarily on providing Internet access to large numbers of people using existing telephone infrastructure. Access is typically provided by wireline telephone systems where such infrastructure is available, but may also be supplied by cellular telephone systems in urban environments. In most very low-income countries, Internet access ranges from difficult to impossible to obtain. Indeed, even conventional telephone services are not widely used. For the year 2001, for example, in India the penetration rate of main telephone lines was about 6%, the average wireless telephone subscription rate was about 0.4%, and the estimated fraction of Internet users was 0.68%, indicating that Internet access was actually slightly more popular than mobile telephone purchases [ITU03]. In some of the world's poorest countries, these numbers fall significantly: in Niger, for example, the cellular subscription rate is about 0.02%, wireline phone rate is 0.19%, and Internet access rates are about 0.11%. Interestingly, we may make two observations: first, it seems apparent that although conventional telephone lines are available in virtually all countries in the world, they are still far from being ubiquitous. Second, we find that for those countries where both Internet and cellular telephones are available, Internet access appears to be more compelling.

As suggested above, in many poor countries neither cellular nor wired telephones are available or affordable. Satellites may serve such areas: either low-earth-orbiting (LEO) or geostationary satellites. Geostationary satellites do not generally provide worldwide coverage, and are typically positioned to provide services to first-world customers. LEO satellite services may be available at much reduced cost with wider world coverage, but are non-interactive and are frequently disconnected [VC03]. Thus, any information access system that depends on a wide-area network in areas of the world lacking significant telephony infrastructure must deal gracefully with the intermittent connectivity properties of LEO satellite systems. In addition, for those locations where some telephony-based infrastructure is available, connection reliability may be low, requiring attention to deal robustly with such (not infrequent) failure.

The three-layer architecture requires networking between the proxies and devices, and between the proxies and data centers. In particular, the proxy may need to deal with two distinct forms of intermittent connectivity. Each form helps to dramatically reduce power and cost at the expense of network latency.

First, it may need to handle unexpected and/or scheduled wide-area connectivity to reach the data centers. Second, it will need to handle the connectivity of asynchronously roaming personal devices that may be operating at low duty cycle. Each of these requirements poses significant networking challenges. For example, many client/server applications are engineered with an inherent assumption that an end-to-end network path exists between client and server, and consider the lack of such connectivity to be an unrecoverable error. In addition, network routing software, responsible for computing shortest paths to network destinations, operates on a routing graph consisting of one large connected component. Most such systems fail to contemplate a network that partitions over time, especially if such a partitioning may be predictable (as in the case of LEOs, for example).

In this project, we propose to adapt, deploy in the field, and continue development of the Delay Tolerant Networking (DTN) Architecture [Fal03] being developed within the Internet Research Task Force's *Delay Tolerant Networking Research Group*, for use with our three-layer architecture. Senior Personnel Kevin Fall is the current chair for the group (see <http://www.dtnrg.org>). Members of the group have developed a design for a DTN software agent capable of handling network message routing over diverse networks (including Internet and non-Internet like networks) where disconnection is considered a comparatively normal occurrence. It operates as a node in a network overlay. Existing DTN components include an application library and a prototype forwarding agent that handles message routing and provides persistent storage in the case of failures. It is capable of handling pre-scheduled connections and reliable message delivery across a network without end-to-end connectivity. Our aggressive use of intermittent networking will require substantial enhancement and evaluation of the current DTN work.

Key Deliverables:

- Novel protocols for applications in the presence of intermittent networking.
- Evaluation of the tradeoffs among latency, cost, power and bandwidth.
- Demonstration of at least 10x improvement in cost and power over cellular networks.
- Demonstration of effective network scheduling (for LEOs and low duty-cycle devices).

LOW-POWER DEVICES AND NETWORKING [RABAAY, BREWER]

From the preceding sections, we can derive the essential constraints and properties of the mobile access devices:

- Manufacturing cost should be as low as possible - this translates into an architecture built around a very small number of simple components.
- Keeping power dissipation as low as possible is essential. The device should be able to operate for a long time (> weeks) on a single battery charge, or should replenish its energy storage by scavenging from external resources such as light and motion.
- The architecture of the device should be modular and scalable. Different applications require devices with very different functionalities in terms of user interface, sensor interfaces, and computational capacity. These requirements naturally lead to a network-centric architecture that consists of a small core centered on communication and storage capabilities. A modular extension capability allows for the addition of computational, user-interface and sensor modules as needed.
- The focus on intermittent networking allows for the radio to be totally powered down for most of the time, preserving precious energy. Approaching a proxy should cause an automatic power up, and trigger a data-interchange session. It should be observed that this intermittent communication strategy requires the availability of ample storage capacity on the device, providing temporary buffer space for data and messages between communication sessions.
- Communication between devices and proxies should be spontaneous and require no set-up or management. In other words, the device should be trivial in its adoption and its maintenance.
- Finally, it would be desirable that the device is upgradeable in its functionality, both by adding hardware modules and through (authenticated) over-the-network software upgrades.

Meeting all the above requirements is absolutely non-trivial and requires an architectural approach that fundamentally departs from traditional portable PC, which is centered on a processor-memory combination. This architecture is useful when computation is truly at the core of the device operation, but brings with it a large inherent overhead. Worse, communication is an add-on, and is not at the core of the device. The Simputer is an example of a device of this class [Iye02].

We propose an architecture that in its bare essence only contains the messaging functionality; this is the combination of communication and storage. Messages to be transmitted are stored in non-volatile memory, waiting for the next communication session. The wireless communication unit is in power-down mode until it detects a nearby proxy. Upon detection, a communication session is initiated, transmitting stored messages and receiving data intended for the device. The average power dissipation of the core configuration can be substantially below 1 mW, enabling extended operation. An energy scavenging power train based on solar power and/or vibration can easily sustain this demand [RWP02]. We project that this basic configuration, when produced in volume, could be manufactured for at most a couple of dollars. We expect this low cost to be enabled by the following technology advances:

- The use of resonant RF-MEMS (such as the Agilent FBAR) enables the design of simple (<5 mm² in silicon) radio modules that are extremely power-efficient and can support 150 kBit/sec [RAK+02].
- The price of high-density non-volatile flash memory is dropping at a very rapid rate.
- Energy scavenging power trains, combining power source, temporary power storage on super-capacitors or rechargeable batteries, and regulators, have been demonstrated [RWP02]. Although the cost currently comes in at \$5-10, this will come down to approximately \$1 over the next 5 years.

The core configuration provides little support for computation. An embedded microcontroller implements the basic storage-communication messaging mechanism, and has some leftover capacity for additional functionality such as data acquisition, sensor interfacing, etc. The provision of more advanced user interfaces (such as speech processing), demanding I/O processing or involved computational functions requires the addition of extra modules. Therefore, the core architecture would embrace a modular bus architecture that supports the insertion of plug-in modules.

We propose to develop a prototype of the core “messenger” in a couple of phases. The first version will be built using commodity components and Bluetooth networking. The second will consist of an off-the-shelf processor with a custom radio module, currently under development at the Berkeley Wireless Research Center [RAK+02]. In a third phase, an integrated system-on-a-chip version would be developed, further reducing costs.

Key deliverables:

- Messenger Core with bus for plug-in modules, with < 1mW power dissipation.
- Development of at least three modules, including a CPU sufficient for speech recognition.
- Support for multiple sensors, via the module interface.
- Demonstration of an energy scavenging power train sufficient for the core.
- Demonstration of ultra low-power intermittent wireless networking.
- Demonstration of variable range wireless networking, from meters to kilometers (with the power tradeoff)
- Demonstration or design for a single-chip device under \$1 if produced in volume.

ULTRA-LOW-COST FLEXIBLE DISPLAYS [SUBRAMANIAN]

The display is often the most expensive item in modern computing appliances [JWS97]. The move to a reel-to-reel fabrication process for display fabrication is the holy grail of much display research, since it will substantially reduce process complexity and cost as a result [Sub02]. An additional advantage of this scheme is that it inherently involves the use of such low-cost flexible substrates as plastic and steel foils, hence enabling lower-cost packaging, greater durability, and unusual form factors.

We will develop a new class of displays by exploiting the fact that the conventional LCD technology is overkill for the target domains; in particular, there is rarely the need for high-resolution high-speed displays. Therefore, we will develop ultra-low-cost displays that will achieve a dramatic reduction in cost by trading off resolution and speed for process simplicity and cost. The entire display will be fabricated using additive reel-to-reel printing technology, using a spin-off of the inkjet circuit technology developed in ECS-0220931 (covered as part of NSF Prior Work). This process eliminates the need for conventional lithography, vacuum processing, and all subtractive steps, which are typically the dominant costs.

The displays will be designed to work with ambient light rather than generating their own light (i.e., they will be reflective displays). This enables the use of cheap steel-foil substrates, which offer excellent mechanical strength and good temperature stability. Reflective displays eliminate the need for a backlight, which will greatly reduce power consumption (displays are historically the largest consumer of power in laptops). In comparison to emissive displays, such as organic LEDs, this technology also offers the advantage of reduced current requirements from the access transistors, since the pixels do not have to deliver the power to generate the light; they are merely capacitively driven to block or transmit the reflected light. This enables the use of a low-performance printed transistor technology.

The pixels themselves will be formed using polymer-dispersed liquid crystals (PDLC) [CNPS99]. A PDLC is a continuous medium of isolated liquid crystals existing in the form of micron-sized droplets dispersed in a polymer binder. Both normally black and normally white modes have been demonstrated. PDLC pixels require no polarizer layers and are formed using polymer materials that can be processed by solution, making them suitable for fabrication using printing. These pixels have high brightness and wide viewing angles; their drawback is that they typically have poor contrast and require higher operating voltages. This is a worthwhile tradeoff, since the low contrast is likely adequate for the text and line-art applications herein, while the higher operating voltage (typically 10V-40V) can be achieved using charge pumps and compatible transistor technology, which we have already developed under ECS-0220931.

Initially, we will develop all the requisite materials and technologies, and will use them to create small pixel arrays as a demonstration vehicle for the power of the technology. Subsequently, we will work to demonstrate a fully functional display with the requisite speed and resolution, all delivered in an ultra-low cost, flexible, and robust form factor. Such a technology could have wide-ranging implications beyond the specific devices developed within this proposal; ultra-low-cost displays are a broadly enabling technology for developing regions, and could be used in a wide variety of low-cost and portable information dissemination appliances with applications extending beyond those targeted in this proposal.

Key Deliverables:

- Production of displays using inkjet technology.
- Integration of displays with final devices.
- Low cost (forecast \$1-5 for a 1/4 VGA display) low-power display technology

USER-INTERFACE AND SPEECH RESEARCH [MANKOFF, CANNY, DEY, WOOTERS]

Although any interface should be built with the specific end user community and application in mind, there are issues we can address in support of the five application domains introduced earlier. A common theme throughout the discussion of the framework and domains is heterogeneity.

Different end users will use devices with varying characteristics in a variety of situations. End users will use different languages to communicate and have varying skills in using IT and abilities in using various input techniques (spoken language, written language, etc.). Devices will have different characteristics including available bandwidth, available services while disconnected, support for mobility, and sensing abilities. Situations will vary with respect to the services available, the task being performed, and the number of people being supported.

With this amount of heterogeneity, the design of a single user interface to support all situations is impossible. However, at a lower layer, common toolkit support can be provided to support a number of user interfaces, customized for different situations. Berkeley and the PIs have a strong history of user-interface toolkit design for speech, gestures, sensor input and other ambiguous forms of input [DSA01, DMAC02, HLLM02, Kle+00, MSA00]. We plan to design a novel interface toolkit that will support a variety of interfaces for different usage situations.

We cannot assume that all users of the systems will be literate, so that spoken language input and output will play a central role in the framework. Of course, reliable recognition of large-vocabulary fluent speech is beyond the state of the art. But for isolated words or phrases from a limited vocabulary, there are cheap, reliable, high-accuracy systems for major languages in widespread use. Similarly, it is straightforward to provide fixed (canned) speech responses in even very inexpensive systems and both of these capabilities have been included in the toolkits described above. The ICSI group that is addressing this task has considerable experience in all aspects of human speech and language processing [Mor+03, SJW+97,

Woo93]. Some initial experiments [Jai02, JR02, Mou02] have shown that speech interfaces can be effective for illiterate and semi-literate users. In fact, one important application is simply to advance literacy and the use of the national language.

The technical challenges lie in making these restricted speech interaction capabilities portable to local dialects without expert intervention. We envision a collection of scripts that can be run by someone with knowledge of a major language and the local dialect, but no programming ability. The speech recognition interface to the system will be deployed with a base set of models for each of the target words or phrases. In situations where adequate language-specific training data are available prior to deployment of the system, these initial speaker-independent (SI) models should work well. However, the initial models may be severely under-trained, in which case the shared infrastructure allows us to collect samples over time, improve the models, and then deploy updates to the models over the network.

There are two approaches that can be used to improve the accuracy of the under-trained models. First, users may be requested to speak examples of some or all of the words/phrases that are used in the speech interface. These example utterances are then used to improve the original set of SI models using standard statistical adaptation techniques. In the second approach, the non-speech UI components are employed to provide a feedback mechanism to the system, which improves the models over time. Thus, the more users speak to and interact with the system, the better the system will recognize the local dialect. Other users and researchers may leverage this learning via the shared infrastructure.

Key Deliverables:

- Novel UI toolkit that supports speech, gestures, and specialized sensors.
- Validate toolkit with applications and two pilot deployments.
- Requirements analysis for the two pilot applications.
- Usability evaluation for the two pilot deployments.
- Training scripts for local dialects usable by non-programmers at the location.
- Demonstration of improved speech recognition via centralized learning over time and over-the-network model updates.
- Demonstration of speech recognition for new dialects with minimal support or cost.

5. PILOT DEPLOYMENT EVALUATION

As stated above, we will execute two pilot deployments in two different regions. For each deployment we will perform two kinds of evaluations (in addition to the evaluation of technology): application success and usability evaluation. We expect to complete multiple rounds of each evaluation.

APPLICATION SUCCESS (EVALUATION OF SOCIAL IMPACT) [RAY, WEBER]

We shall conduct an evaluation of the social and economic impacts of our proposed pilot deployment in two stages of this research program. The first will be at the time of deployment itself, and the second towards the end of the final year. In each case, we shall compare the pre-deployment and post-deployment:

- Overall characteristics of the chosen communities, focusing on those we expect to be most changed by the specific applications (measured by median incomes, radius of market trades, access to healthcare, literacy levels for men and women);
- The specific characteristics of the lower socio-economic strata within these communities (e.g. incomes, market participation or education levels); and
- The region-wide levels of income, health status etc, beyond the pilot locations (these data sources will be, of necessity, secondary – but will help us to separate out the effects of our interventions from any concurrent regional changes).

This comparison will allow the team to assess:

- The overall *and* relative welfare gains (if any) from the new technologies, and
- Any unintended consequences of the technology deployments.

USABILITY EVALUATION [MANKOFF, CANNY, DEY]

Our usability evaluation will be closely tied to measures of application success. We will use a combination of interviews, focus groups, and field experiments to learn about the usability of the applications we develop. We choose these qualitative techniques because they will give us the most

complete understanding of the use of our applications in the field. Long before deployment, we will have worked out basic usability issues through early stage low-cost and lab based evaluations.

The methods we plan to use in the field will help us to develop an understanding of how the applications succeed or fail at the level of individuals or small groups. We will tie these techniques into an iterative design process in which we work to address the specific application issues that we believe may be affecting application success, based on our observations and other data.

6. INTEGRATION OF RESEARCH AND EDUCATION

CS160 User Interface Design: Berkeley has a long tradition of human-computer interaction research that is guided by a deep understanding of the needs of the people being supported [IML03, NM00, KL00], and we plan to continue in that vein in our work here. As we identify specific target communities needing support in each of our application areas, we will work closely with them to identify their needs, conducting on-site interviews and observations that will guide the technologies we design.

Along those lines, we plan to focus our undergraduate human-computer interaction class (CS160) on those application areas, encouraging them to reach out to local poor, rural communities as they develop their group projects. Students are required to work very closely with their customers, interviewing them and testing the effectiveness of their software with customers three different times over the course of the semester. We plan to select the most successful students for a one-month trip to one of the actual developing regions being supported by our research efforts, where they will conduct another round of the development cycle emphasized in the class.

In the past few semesters, CS160 has been very successful in repeatedly focusing on the social needs of local community members. Last spring, the class worked with the Institute for Women and Technology to target local women with disabilities [Qui02]. This fall, students were required to help out in a local soup kitchen and focus their projects on the needs of local non-profit organizations. Students' comments on the service work were very positive.

Graduate Education: We propose to teach two different advanced research seminars on technology infrastructure for developing regions. As with our other graduate courses, we would also welcome advanced undergraduates. To be offered in Fall 2003, the first is a multi-disciplinary course that looks at the target application domains and regions. We expect this course to provide a unique educational opportunity, and also to help define the two primary deployment projects. These seminars would consist of paper reading, guest lectures, and a course project that involves a proposed infrastructure-supported solution for a specific application in a specific region. In addition, the existing graduate "field methods" class, taught by Senior Personnel Isha Ray, teaches the key social science evaluation methods that we will use for our applications, and real-life issues from the pilot deployments will be used to enhance the class.

The second course will be more technical, and will use class projects to find innovative ways to co-design applications, devices and infrastructure. These seminars would also consist of weekly research paper reading, guest lectures, and a course project. Students would be expected to develop working prototypes of co-designed servers, proxies and devices for a specific application. Similar course projects will also be allowed for the core graduate CS classes.

Finally, the investigators have a strong record of working with both graduate and undergraduate researchers. We intend to continue student involvement with the project outlined here by supporting up to 12 graduate students and 2-4 undergraduates. The PIs have a track record of involving underrepresented groups in their research and they will continue to do so with this work.

Support for Diversity: In addition, we are sponsoring eight summer internships for undergraduate engineering students from underrepresented groups, via the successful SUPERB NSF REU program. Each of the eight will be supervised by one of the PIs for an immersive research apprenticeship. We believe IT for Development applications will affirm students' motivation for graduate study and strengthen their qualifications, while demonstrating the potential of IT for positive social impact.

7. SUMMARY

We hypothesize that by directly attacking the key technical issues of developing regions, we can far outperform the traditional approach of "trickling down" IT that was designed for affluent regions. There

are several technical innovations in this approach. First, we focus on low-cost low-power devices that move away from the idea of a stripped-down PC, focusing instead on communication and dependence on shared infrastructure. Second, we exploit intermittent networking at all levels because it is an order-of-magnitude better in both cost and power consumption over traditional continuous real-time networks; it is also sufficient for a very wide range developing region applications. Third, we develop novel core technologies that are unlikely to arrive on their own, including ultra-low-cost displays (but lower fidelity), support for groups with low levels of literacy and a wide range of languages or dialects, and low-power intermittent wireless communication.

We aim to enable a very broad range of ICT applications for developing regions, and will validate the framework through two pilot deployments. Each deployment will be evaluated for application success and usability, and we expect in the process to address two important ICT applications. We have a remarkable set of partners to provide guidance and deployment help, including Intel, HP, UNDP, Markle, IIT Delhi, CITRIS and Grameen Bank. Finally, we will seek additional funding for refining all aspects of this activity, and believe that some of the innovations will help the first world as well.

8. MANAGEMENT PLAN

PI Brewer will serve as the project director and will manage the project. Although key research and administrative decisions will be made in a collaborative manner, Professor Brewer will serve as the contact point for the NSF and our industrial collaborators, and have responsibility for managing the administrative personnel who will handle the project budgets and spending. We will continue our current mode of management, which is based on weekly in-person meetings, augmented by conference calls. All of the PIs, Senior Personnel, students and staff are local to Berkeley.

The EECS department at UC Berkeley has had considerable success with twice a year off-site project retreats. These meetings are sponsored by industrial funding and offer a chance for the students to present their work in detail while getting feedback from the entire team as well as invited industrial and academic guests. They also offer the project team a good place to schedule the following six months of activity.

It is also our expectation to be able to use our own project to facilitate management and communication. We will build a local version of the architecture on campus to enable development and experimentation. As a side effect, team members will be able use our low-power communication devices as research and communication tool starting sometime in year 2.

Conferences: To promote multi-disciplinary research and collaboration, we plan to hold at least two conferences on technology for developing regions. The first one is “Enabling Digital Technologies in Emerging Markets” in Summer 2003, which is actually before this project would officially begin and is therefore funded separately. We expect to gain insight from these conferences into the best applications and regions for our two deployments.

Managing facilities: We will benefit from the use of existing staff people (separately funded) to maintain and operate the data center that we will use for our testbeds and deployments. We expect successful deployments to be transferred to commercial data centers at the end of the project. There is also a staff of several people on other funding to maintain the development resources, manage the network, and help with software upgrades.

Deployment Management: A critical success factor for this project is the delivery of real testbeds and deployments. PI Brewer has a long track record of developing and operating live 24x7 services both in industry and for research projects, particularly the NOW project, which had 50,000 users per day at one point. In addition, the budget for years 2-4 includes funding for a staff person to manage live deployments (part time in year 2, full time for the final two years). The person will manage upgrades to the infrastructure, manage the source code repository and its remote developers (including for example researchers at HP Labs in Bangalore, India), and act as the liaison for partners managing remote deployments. Once we get closer to deployments, we expect to start weekly conference calls that include the remote partners. In addition, we have requested a significant travel budget to facilitate the key aspects of deployment: requirements gathering at the beginning, deployment facilitation in the middle, and some end-of-project evaluation at the end.

Outside guidance (CITRIS): This project will be part of and sponsored by CITRIS, the Center for Information Technology Research in the Interest of Society, a new multi-campus research institute

recently founded at UC Berkeley, UC Davis, UC Santa Cruz and UC Merced (www.citris.berkeley.edu). The purpose of CITRIS is to support collaborative, IT-focused research to find solutions to grand-challenge social and commercial problems affecting the quality of life of individuals and organizations. The Technical Advisory Board of the center is extremely interested in this project and will continue to provide advice and guidance. The board is comprised of distinguished researchers, VP's of engineering and CTOs of several major IT corporations, including BroadVision, Ericsson, Hewlett-Packard, IBM, Intel, Marvell, Microsoft, Nortel Networks, ST Semiconductors, and Sun Microsystems. CITRIS will also commit up to 5,000 ASF in the new CITRIS II building, presently under planning and construction, to house the key collaborative elements of the project. We also anticipate the very large server farm, housed in a secure and independently powered sub-basement of CITRIS, will be used to support this project.

Financial Management. Due to the complexity and global scope, top administrative staff will handle the contracting end of the project. A portion of a staff person is provided in the budget. She will provide specific support functions for contract/grant compliance, including: post-award grant management, financial analysis, projections, record keeping, billing coordination, procurement, travel, reimbursement and travel voucher preparation, receiving and processing deliveries, monitoring equipment inventory, and payroll (initial hiring, processing, etc.). These services are not included in the campus indirect cost rate, nor do central campus units provide them.

9. PARTNERS

CITRIS will also contribute 10 percent of Thomas Kalil's time to help manage and expand the project's relationships with external partners. Kalil serves as the Special Assistant to the Chancellor for Science and Technology.

Intel: Intel has long history of support for Berkeley, including donating much of the equipment that will be used for the data-center layer. Intel will support 100% of two of the Senior Personnel, Kevin Fall and Anind Dey. In addition, Intel is developing a small X-Scale based computer that could form a low-cost low-power proxy; Intel will donate at least 20 of these proxies. Finally, Intel will provide deployment guidance via its "People and Practices" research group, which includes anthropologists and other social scientists that are deeply familiar with human behavior and the use of IT in developing regions.

Grameen: Muhammad Yunus is the founder of Grameen Bank, which has provided "micro credit" to 2.4 million poor women in rural Bangladesh. Grameen is widely viewed as one of the most important global models for successfully extending credit to the poor. Dr. Yunus has also been involved in the creation of a number of other enterprises, such as Grameen Telecom. This company provides a GSM phone to "phone ladies" in rural villages, who earn money by renting it out to as many as 70 other people in the village. Grameen is interested in serving as a possible testbed for the deployment of Berkeley technologies.

Hewlett Packard: HP is one of the founding corporate members of CITRIS, and is committed to investing \$6 million to support CITRIS research during the next four years. Recently, HP has launched an "e-inclusion" initiative to close the gap between technology-empowered and technology-excluded communities. HP is interested in collaborating with Berkeley on this project in three ways. First, they are interested in fostering joint research projects with HP Labs India, which is focused on creating ICT innovations that address the needs of emerging markets such as India. Second, they are willing to use their two "i-Communities" as possible testbeds for Berkeley technologies. These "i-Communities" will demonstrate the potential of technology to help improve literacy, job creation and income, government services, education, and healthcare. Third, they are willing to provide \$50K in seed funding, to be used at the discretion of the PI, Eric Brewer.

Markle Foundation: The Markle Foundation works to realize the potential of information technology to improve people's lives. Markle has served on the U.N. ICT Task Force and as the U.S. non-profit representative to the G-8 Digital Opportunity Task Force. Markle is supporting a number of initiatives designed to expand digital opportunities in developing countries. For example, Markle has funded the creation of an ICT Working Group within InterAction, an association of 160 U.S. nonprofit organizations engaged in international relief and development work. This working group is helping U.S. non-profit organizations (and their developing country NGOs) share "best practices" on innovative applications of

IT for development. InterAction's members have first-hand experience in using IT in some of the poorest and most remote places on the planet, which includes coping with extreme heat, sand storms, shortages of trained personnel, and lack of power and network connectivity. Through InterAction, UC Berkeley will gain valuable insight into the needs and challenges in developing countries.

India Institute of Technology Delhi: IIT Delhi has recently launched major, government-funded research projects in low-cost computing and sensor networks. They are interested in partnering with Berkeley to (1) team teach a short course on IT for developing countries that would be offered in both India and Berkeley; (2) host visiting faculty and students from Berkeley; and (3) share their analysis on application requirements for low-cost computing in developing countries.

United Nations Development Programme: UNDP leads the U.N.'s efforts to make new technologies for human development, and has active programs in 166 countries. The Administrator for UNDP has expressed interest in providing input on the range of IT applications that are important to support for development purposes, and the economic and technical challenges associated with offering affordable access to disadvantaged communities.

We anticipate adding additional partners as the project moves forward and we identify interested parties.

REFERENCES

- [BB98] E. S. Berner and M. J. Ball, editors. *Clinical Decision Support Systems: Theory and Practice*, Springer-Verlag, 1998.
- [Bre01] E. Brewer. "Lessons from Giant-Scale Services." *IEEE Internet Computing* 5(4): 46-55. 2001. <http://www.cs.berkeley.edu/~brewer/papers/GiantScale.pdf>
- [Bre+02] E Brewer, N. Borisov, M. Chen, R. von Behren, M. Welsh, D. Culler, J. MacDonald, J. Lau, and S. D. Gribble. "Ninja: A Framework for Network Services." *Proceedings of the 2002 Usenix Technical Conference*, June 2002, Monterey, CA, <http://www.cs.berkeley.edu/~brewer/papers/Ninja-Usenix2002.pdf>
- [Byr+00] W. Byrne et al., (2000), "Towards Language Independent Acoustic Modeling," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Istanbul. II-1029.
- [Can02] J. Canny, "Collaborative Filtering with Privacy." *IEEE Conference on Security and Privacy*, Oakland CA, May 2002.
- [Cas01] M. Castells. "The Internet Galaxy: Reflections on the Internet, Business, and Society." Oxford University Press, 2001.
- [CDT02] Center for Democracy and Technology and InfoDev. *The E-Government Handbook for Developing Countries*. November 2002. <http://www.cdt.org/egov/handbook/2002-11-14egovhandbook.pdf>
- [CNPS99] H. Cornelissen, J. Neijzen, F. Paulissen, L. Schlangen, "Reflective direct-view LCDs using PDLC and dielectric reflectors." *Journal of the Society for Information Display*, Vol. 7, No. 1, 1999, pp. 37-41.
- [Coh01] N. Cohen. "Rural Connectivity: Grameen Telecom's Village Phones." *World Resources Institute*, June 2001 (online at <http://www.digitaldividend.org>).
- [DMAC02] A. K. Dey, J. Mankoff, J., G. D. Abowd, and S. Carter. "Distributed Mediation of ambiguous context in aware environments." In *Proceedings of UIST 2002*, pp. 121-130.
- [DOI01] "Creating a Development Dynamic: Final Report of the Digital Opportunity Initiative," July 2001. Available at: <http://www.opt-init.org/framework/pages/contents.html> [Fal03] K. Fall, "A Delay Tolerant Network Architecture for Challenged Internets", submitted for publication, Feb. 2003
- [DSA01] A. Dey, D. Salber, G. D. Abowd, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications," anchor article of a special issue on Context-Aware Computing, *Human-Computer Interaction (HCI) Journal*, Vol. 16(2-4), 2001, pp. 97-166.
- [DR02] J.B. Downing and I. Ray. "Safe Drinking Water for Everyone", *mimeo*, Energy and Resources Group, University of California at Berkeley, 2002, pp. 25-26.
- [FGC+97] A. Fox, S. Gribble, Y. Chawathe, E. Brewer and P. Gauthier. "Cluster-Based Scalable Network Services." *Proceedings of SOSP '97*, St. Malo, France, October 1997. <http://www.cs.berkeley.edu/~brewer/papers/TACC-sosp.pdf>
- [FGCB98] A. Fox, S. Gribble, Y. Chawathe and E. A. Brewer. "Adapting to Network and Client Variation Using Active Proxies: Lessons and Perspectives." *IEEE Personal Communications*, Special Issue on Adaptation, August 1998. <http://www.cs.berkeley.edu/~brewer/papers/proxy-IEEE.pdf>
- [GLvB+03] D. Gay, P. Levis, R. von Behren, M. Welsh, E. Brewer and D. Culler. "The nesC Language: A Holistic Approach to Networked Embedded Systems." To appear, *Proceedings of PLDI 2003*. <http://www.cs.berkeley.edu/~brewer/papers/nesC-pldi.pdf>
- [GTR02] G. Kirkman, editor. "The Global Information Technology Report 2001-2002: Readiness for the Networked World," *Oxford University Press*, March 2002, (selected online chapters available at http://www.cid.harvard.edu/cr/gitrr_030202.html)

- [HLLM02] J. I. Hong, J. A. Landay, A. C. Long, and J. C. Mankoff. "Sketch Recognizers from the End-User's, the Designer's, and the Programmer's Perspective." In *Proceedings of AAAI 2002 Spring Symposium (Sketch Understanding Workshop)*.
- [HLS02] D. C. Huang, F. Liao, and V. Subramanian. "Inkjetted gold conductors for electronics on plastics", presented at *Materials Research Society Spring 2002 Meeting*.
- [IML03] M. Ivory, J. Mankoff and A. Le. "Using Automated Tools to Improve Web Site Usage by Users with Diverse Abilities". In *Information Technology and Society*. To Appear.
- [Int03] Intel X-Scale Processor Information. <http://www.intel.com/design/pca/applicationsprocessors>
- [Iye01] S. Iyengar. "India to Compute on the Cheap." *Wired News*, June 28, 2001.
- [ITU03] Statistics on wired telephone, cellular, and Internet use are available at <http://www.itu.int/ITU-D/ict/statistics>
- [Jai02] S. Jain. "A Reading Tool for the Illiterate." Proc. of the Second International Conference on Open Collaborative Design for Sustainable Innovation, Bangalore, India, Dec. 2002.
- [JR02] A. Joshi and S. Rasal. "Interaction Paradigm for Semi-literates and Illiterates." *Proc. of the Second International Conference on Open Collaborative Design for Sustainable Innovation*, Bangalore, India, Dec. 2002.
- [JSW+00] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister. "System architecture directions for network sensors." *Proceedings of ASPLOS 2000*, Nov. 2000.
- [JWS97] S. Jurichich, S. Wood, and K. Saraswat, "Cost modeling of low-temperature large-area polysilicon thin film transistor liquid crystal display manufacturing", *Proceedings of SPIE - the International Society for Optical Engineering*, vol. 3014, 1997, pp. 160-165.
- [KL00] H. M. Kamel and J. A. Landay, "A Study of Blind Drawing Practice: Creating Graphical Information Without the Visual Channel." In *Proceedings of the Third International ACM SIGCAPH Conference on Assistive Technologies: ASSETS 2000*, Washington, DC, November 13-15, 2000, pp. 34-41.
- [Kle+00] S. R. Klemmer, A. K. Sinha, J. Chen, J. A. Landay, N. Aboobaker, and A. Wang. "SUEDE: A Wizard of Oz Prototyping Tool for Speech User Interfaces." In *Proceedings of UIST 2000*, pp. 1-10.
- [Koh00] Vinod Khosla, comments at the *Creating Digital Dividends Conference*, October 16, 2000 in Seattle, WA
- [Koh01] J. Kohler. (2001), "Multilingual Phone Models for Vocabulary-Independent Speech Recognition Tasks," *Speech Communication*, vol. 35, pp. 21-30.
- [LC02] P. Levis and D. Culler. "Mate – a Virtual Machine for Tiny Networked Sensors." *Proceedings of ASPLOS 2002*. Oct. 2002.
- [LC03] M. C. Linn and M. J. Clancy. "Overview of the UC-WISE Project" <http://buffy.eecs.berkeley.edu/~clancy/UCWISE.overview.html>
- [MHA00] J. Mankoff, S. E. Hudson, and G. D. Abowd. "Providing Integrated Toolkit-Level Support for Ambiguity in Recognition-Based Interfaces." In *Proceedings of CHI 2000*, pp. 368-375.
- [Mei95] G. M. Meier, editor. "Leading issues in economic development". New York, NY. Oxford University Press, 1995.
- [Mor+03] N. Morgan, D. Baron, S. Bhagat, H. Carvey, R. Dhillon, J. Edwards, D. Gelbart, A. Janin, A. Krupski, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters. 2003. "Meetings about meetings: research at ICSI on speech in multiparty conversations", ICASSP-2003, Hong Kong, April 2003.
- [Mou02] A. Mougald et al. "A Toolkit for Building Accessible User-Interfaces." Proc. of the Second International Conference on Open Collaborative Design for Sustainable Innovation, Bangalore, India, Dec. 2002.

- [MPS+02] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, J. Anderson. "Wireless Sensor Networks for Habitat Monitoring." *2002 ACM International Workshop on Wireless Sensor Networks and Applications*. September 28, 2002.
- [NEES] Network for Earthquake Engineering Simulation. <http://nees.berkeley.edu/>
- [NL00] M. W. Newman and J. A. Landay. "Sitemaps, Storyboards, and Specifications: A Sketch of Web Site Design Practice." In *Proceedings of Designing Interactive Systems: DIS 2000*. pp. 263-274, August 17-19 2000.
- [Pat02] D. Patterson. "E-Mexico Marks Country's Entrance into the Digital Age." *Government Technology*, April 19, 2002. <http://www.govtech.net/news/news.phtml?docid=2002.04.19-3030000000006201>
- [PC98] E. Paulos and J. Canny. "Prop: Personal roving presence". *ACM Conference on Human-Computer Interaction (CHI)*, 1998.
- [Pix03] Pixim. Digital Pixel Technology. http://www.pixim.com/pt/pt_dps.htm
- [Pri82] J. Price Gittinger. "Economic analysis of agricultural projects", (2nd ed.). Baltimore: Johns Hopkins University Press, 1982
- [Qui02] F. Quintero. "Disability Brainstorm: Computer science students ponder innovations." *Berkeleyan*, February 14, 2002. http://www.berkeley.edu/news/berkeleyan/2002/02/14_disab.html
- [RAK+02] J. Rabaey, J. Ammer, T. Karalar, S. Li, B. Otis, M. Sheets, and T. Tuan, "[PicoRadios for Wireless Sensor Networks: The Next Challenge in Ultra-Low-Power Design](#)" *Proceedings of the International Solid-State Circuits Conference*, San Francisco, CA, February 3-7, 2002.
- [RWP02] S. Roundy, P. K. Wright and K. S. J. Pister, "Micro-electrostatic vibration-to-electricity converters", *Proceedings of ASME IMECE*, in New Orleans, LA, Nov. 17-22, 2002.
- [SC97] J. D. Slotta and A. Cuthbert. "Applying a framework of conceptual change in the science classroom: Studies of knowledge integration." *Proceedings of the annual meeting of the Cognitive Science Society*, 1997
- [SIL02] SIL International (2002), "Ethnologue: Languages of the World, 14th Edition", www.ethnologue.com
- [Sub02] V. Subramanian. "Prospects for organic semiconductor based devices and circuits: applications, performance, and reliability considerations." *2002 Topical Research Conference on Reliability*, Oct. 2002.
- [SW01] T. Schultz and A. Waibel. "Language-Independent and Language-Adaptive Acoustic Modeling for Speech Recognition." *Speech Communication*, vol. 35, nos. 1-2. 2001.
- [SWJ+97] J. M. Sackier, M.D., C. Wooters, L. Jacobs, M.D., A. Halverson, M.D., D. Uecker, Y. Wang. 1997. "Voice Activation of a Surgical Robotic Assistant." *The American Journal of Surgery*. Oct;174(4):406-9
- [UN01] United Nations Development Programme. "Human Development Report 2001: Making new technologies work for human development." 2001. <http://hdr.undp.org/reports/global/2001/en/>
- [VC03] See the model proposed by the Vita-Connect program using LEOs at <http://www.vita-connect.org/brochure.html>
- [WB01] *World Bank Development Report*. "Building Institutions for Markets." Oxford University Press, 2001.
- [Woo93] C. Wooters, 1993. "Lexical Modeling in a Speaker-Independent Speech Understanding System." Ph.D. Dissertation, University of California at Berkeley. Berkeley, CA. 1993.
- [WRI03] World Resource Institute. Digital Dividend Project, <http://wriws1.digitaldividend.org/wri/app/index.jsp>