Travels in Silicon

Reflections on the CGEY's Center for Business Innovation's February 2001 Untethered Devices Rave

By Geoff Cohen and Mukul Kanabar

June 2001

How will these tools change the ways in which we use the network to communicate, work, or live?

Travels in Silicon

Reflections on the CGE&Y Center for Business Innovation's February 2001 Untethered Devices Rave

"A stone, even the most roughly shaped pebble, is the culmination of an intricate network of mental and physical processes. There must be an intelligent brain that can see the need for a tool, and also see the finished tool's shape within the shapeless starting lump of rock. The brain must plan how to fashion the rock, how to hold it, and where to strike it, step by step, to achieve the finished article."

-Robin McKie, The Dawn of Man

About two million years ago, some pretty smart apes figured out that tools would help them hunt. They probably didn't expect it, but it was also a key step in their evolution, and forever changed their relationship with the environment. Some time has passed, and again we find ourselves considering how to build some tools. This time, the tools are small, portable, untethered, networked devices: perhaps not as good for bringing down an antelope as those small bits of stone, but with the same potential for transforming the ways in which we touch the world.

With the era of cheap, ubiquitous wireless connectivity almost here, even the tiniest of devices can interact with a network. Traditionally, these devices included mobile phones, pagers, electronic organizers, and notebook computers. But networked devices are evolving into tools that allow us use the network not just to speak to others, but to access data and services, interact with applications, and take action. As the interface between the person and the network, there is an enormous burden on mobile devices to allow us to communicate our needs and desires to the network, and participate fully in the network without losing control.

How will these tools change the ways in which we use the network to communicate, work, or live? What things will we do differently? What new things will we do? Everyone seems to agree that networked devices have almost unlimited potential to enable new applications, to improve people's lives, and yes, to generate enormous profits. And yet despite enormous investments of money, time, and attention over the last few years, they don't seem near to fulfilling that potential.

On February 28, 2001, thirty people gathered at the Cap Gemini Ernst & Young Center for Business Innovation to discuss the future of the networked device. The gathering includ-

...mobile applications have a much smaller margin for error than traditional desktop applications.

ed Internet pioneers, university researchers, venture capitalists, consultants, and designers. This paper draws upon and reflects that conversation, which revolved around three key challenges facing mobile, networked devices: most current applications are poorly suited to the unique needs of mobility; mobile devices typically ignore the importance of user experience; and top-down, centrally-managed wireless networks inhibit innovation. Only once these challenges are met and overcome by wireless providers will mobile devices be able to reach their potential.

The Challenges of Mobility

Mobile applications are fundamentally more challenging to construct than traditional desk-top or web-based applications, and not for merely technological reasons (although indeed they pose significant technical challenge). First, by definition, a user employs mobile devices in a wide variety of locations, contexts, and activities. At different times, a device performs different tasks, with different resources available, and with differing goals and standards for success.

Second, mobile devices have inherently limited capabilities. For both technical and economic reasons, mobile devices have smaller screens, slower processors, and slower bandwidth links, compared to the equivalent fixed desktop computer. Most critically, such devices have a limited power supply, a factor that software designers have had the luxury to ignore for the last few decades. As a result, our programming methodologies, operating systems, and development environments are ill-suited to the construction of power-efficient or even power-aware applications.

Finally, to make matters worse, mobile applications have a much smaller margin for error than traditional desktop applications. When mobile, humans necessarily have their attention diverted. They may be driving, in an unfamiliar location, or simply performing some other task. Mobile devices that demand attention or excessive care and feeding (such as the endless rebooting that desktop users are resigned to) will simply not be accepted. Similarly, the design of the user interface must meet a high standard

if it is to be usable on a one-inch screen; there simply isn't any room to get it wrong.

Meeting these challenges will require cooperation and innovation from service providers, device manufacturers, and application builders. The fact is that we simply don't yet know how to build these applications well. But there are some guidelines we can learn from early successful mobile models. Mobile applications must have a number of characteristics. They must meet compelling needs. The tools must match the task. Finally, applications must gracefully adapt to changing circumstances.

Meet Compelling Needs

On one hand, mobile devices are convenient, being always close-to-hand and not requiring a wait to boot up or connect to the network. However, when traveling, people have other things on their mind, and the limitations of the device can make people unwilling to spend unnecessary time using them. Mobile applications must surpass these barriers by providing truly compelling applications. They can do this bymaking a task easier to accomplish, by letting users do it more quickly, effectively, or conveniently; or they should enable users to accomplish new forms of tasks previously impossible.

This is why "mobile commerce" (buying things over the Internet using a mobile phone) seems so uncompelling: it is harder to do than using a PC, and it does not let us do much that is new. Some vendors hope to couple m-commerce to locality, letting users receive coupons (or worse, advertisements) for close-by retail outlets. This plan seems dubious at best.

Personal connectivity — communicating with another live human who's in a different place — was of course the original mobile application. Making phones mobile changed the patterns of socializing and work, once connecting with people became simpler and available more of the time. First, you are now reachable: you can be shopping for shoes but your clients can still find you. Second, you can easily connect with others, even when you're away from the home or office. Mobile phones were successful precisely

Mainstream people don't use mobile devices because they enjoy using devices....they do it to accomplish tasks.

because they made a compelling application — talking to other people — easier in some circumstances and possible in previously impossible ways.

But the applications of networked devices go much further than just talking to humans; they can access data, interface with other devices, support decision-making, and provide wayfinding. Compelling applications for these sorts of tasks generally take two forms. The first involves issues around time: applications that are time-sensitive, must be performed at specific times, or have some sense of immediacy or impulse. For example, commercial bankers must often respond within minutes of changes in interest rates. For this, the timeliness of their response is more important than where they perform it. The second set of issues revolves around space. Applications that arise when people are away from their office, in new or different places, or in transit. These applications also allow people to multitask, or fill periods of time that otherwise would be wasted.

Networked devices can also be used as a focus of control. They can change control in two ways: enablement shifts control into a user's hands, while automation shifts it away. In different circumstances each can be appropriate and helpful. Of course, even automation should never rob the user of the ability to make decisions; rather, it should shift the focus of the user's work away from repetitive tasks to higher level decision-making.

Match the Tool to the Task

Mainstream people don't use mobile devices because they enjoy using devices (although such a population clearly exists), they do it to accomplish tasks. And yet too often, operating the device gets in the way of using it. This ranges from endless rebooting to needing to know to which bandwidth range a specific Bluetooth device uses. Most people don't care, and they certainly don't read the instructions to find out.

This preponderance of operation instead of use is a function of many issues. First, good interface design is fundamentally hard, and requires multiple iterations of watching how people use an interface, learning from that experience, and redesigning. Second, there's often a poor match between these devices and

the kinds of tasks they purport to offer, as anyone who's entered a text message with a numeric keypad can attest. Finally, for many of these tasks, we lack the proper abstractions. For example, if I want to tell my phone to screen calls that aren't important, I simply don't have a language that I can speak to the device to tell it how to do that. It's hard enough to tell a human which calls to screen, and humans (mostly) have far better judgment.

Things are not as dire for interface designers as it might seem, however. People are willing to learn how to operate difficult devices, if the use is compelling enough: for most people, driving a car requires classes and hours of practice. And teens seem more comfortable with using networked devices than adults; perhaps older people's standards for ease of use and appropriate interfaces will be irrelevant someday.

Much has been made of convergence, the trend of combining more pieces of functionality into the same box. It's an attractive offer: who wouldn't want to carry just one convenient tool that did everything? But such devices are likely to be less than the sum of their parts. A number of vendors have brought PalmOS-enabled mobile phones to market; as yet, none have succeeded. Fundamentally, the personal organizer and the phone have different sets of requirements, both technologically and in form factor. To combine the two is to compromise on both at higher cost; this leads to collisions like holding a visual interface to your ear. Given sufficient technology, such compromises may become acceptable. Yet it is instructive to look into the kitchen, where no "do-it-all" tool has replaced blenders, food processors, mixers, and bread machines. People keep specialized tools for many reasons: if those tools do the job better, if they are cheaper than a hybrid tool, or if the hybrid tool is liable to need repairs or replacement, for example. While the demands of mobility and convenience will always push for more convergence, these desires for good tools will always push back.

Adapt to New Contexts

By their nature, mobile devices encounter a variety of places, times, and settings. Networked devices must cleanly adapt to their current context, offering different

As networked devices become more sophisticated and do more things, they twine themselves more deeply into our lives. This makes it all the more critical that they enhance our experience, rather than detracting from it.

options, making differentchoices, and acting differently depending on what is around them. While software in general and devices in particular have made much progress in customizing to specific users, mobile applications must go further.

There are three main components of context: the location, the social setting, and the activity. Different places can offer very different problems and resources to a person. Many services, including network nodes, retail outlets, bathrooms, printers, and offices, are inherently localized. Social settings define what sorts of behavior are appropriate. Most challengingly, the activity that a user is engaged in at a single moment will have an enormous impact on how much attention they have, whether interruptions are welcome or not, and what sorts of tasks they are likely to do next.

Some places may be completely foreign: a never-before visited airport, city, or even culture. In these cases we may look to our devices for help wayfinding, locating these services, and more deeply offering advice and guidance. For example, stepping offa plane in an airport in a new country, a user may not even know to ask for directions to the agricultural goods inspection station (I didn't!). A useful device not only responds to queries, it offers timely and tailored advice.

Other places are more familiar: our kitchen, our study, our nursery. While in the past the industry has focused on the study/home office and the living room as the beachheads for new technology, there is a sense that the kitchen (an important center of social activity as well as a focus for concerns over health and food safety), the baby's room, and the basement may be even more important.

Sometimes we're in a place where we have to be, but aren't otherwise occupied, such as waiting in a doctor's office; sometimes we're in a place where we want to be, like the movies. Sometimes we can be interrupted, and sometimes we're with other people and should consider their feelings on the matter. Different interruptions will have different priorities, and it is the proper matching of interruption and activity that has yet eluded communication devices. Of course, I always want to be inter-

rupted in cases involving the health or safety of my family, and never want to be interrupted by telemarketers, but beyond that, it is difficult to express. Even if we could specify the desired behavior, what mechanisms can we build that allow such a world?

Experience

As networked devices become more sophisticated and do more things, they twine themselves more deeply into our lives. This makes it all the more critical that they enhance our experience, rather than detracting from it. For example, the difference between being reachable and being able to make calls is critical. One is an interruption, even if a welcome one, while the other is at your initiation. This tension has created much of the resentment we have over these little phones that are our masters. We welcome the increased ability to reach others they bring us, but we wish we had some control over when they could reach us.

As anyone who's seen a movie in a theater in the last two years knows, calls can be incredibly disruptive and annoying, both for people carrying the phone and everyone around them. There are a number of factors that drive this sense of disruption. The per-minute pricing model of most wireless phones certainly makes us more sensitive to non-critical calls. More inherently, the nature of mobility means that we're likely somewhere where we're explicitly trying to get away: a movie theater, an important meeting, on vacation. We aren't necessarily willing to cut the communications channel off completely -- after all, it might be important. But we would certainly appreciate the ability to control who has access to that channel.

The killer application in this space might, ironically enough, be isolation, not connectivity (giving credence to the belief that every problem was a solution once). It's easy to turn off the phone, but what we want is a careful, selective filtering of who gets to interrupt us. Unfortunately, this is harder than it appears. Only the best human receptionists always know which calls to let through at which times, and that often takes months (if not years) of learning. While some progress has been made in automated assistants, this ability to make context-sensitive judgment calls is currently beyond the ability of even laboratory artificial intelligence. This isn't

It's easy to turn off the phone, but what we want is a careful, selective filtering of who gets to interrupt us.

entirely the fault of the AI researchers; we don't even have a language to describe to humans our preferences for the balance between isolation and reachability.

In a broader sense than just calls, mobile devices are currently quite contemptuous of the value of our attention. Many observers have pointed out that human attention is the last scarce resource; we have exhausted our attention budget on advertisements, operating electronic devices, and periodically checking email, not to mention actually getting work done or not crashing the car. New mobile applications will at least have to be attention-neutral; that is, not require any more diversion of our attention than they return in productivity. Better yet would be applications that solve these problems for us, by filtering out interruptions, by silently automating tasks, or by assisting us with the information we need, just when we need it.

Networked devices also have the potential to deepen, create, or capture experience. The audio guides available at most museums are devices that augment experiences. At best, these guides help deepen and enrich the experience of walking through the exhibit. The danger is that they lock visitors into a preset path, or otherwise interfere with their freedom. Devices such as music players or game machines provide experience. As well, they can capture experience, by taking pictures, movies, or recording sounds. Adding the capability to interact with others through the network offers possibilities for real innovation in these traditional tasks.

Finally, just carrying a device can be part of a social experience, as a fashion accessory or as a sign of status or importance. How many people buy oh-so-fashionable devices to show them off to friends? The fact is that (some) people are fashion-conscious; people buy well-designed devices both because people value good design in itself, and because shiny things are a good way to impress others. In a lot of ways, a mobile phone is equivalent to a car in the 1950s: a way to demonstrate your taste, your wealth, and your importance. I've had colleagues turn away in the midst of a sentence to take a call, holding up their hand; even as they were speaking into the phone, they were sending me a message, too.

The Loneliness of the Long-Distance Network

In the milestone work "The Mathematical Theory Of Communication," Claude Shannon wrote in 1948:

> "The [communication] system must be designed to operate for each possible [message], not just the one which will actually be chosen since this is unknown at the time of design."

The "end-to-end" design principle described in a seminal 1984 paper by MIT researchers Jerome Saltzer, David Reed, and David Clark, applies this principle more broadly to today's networks, and tells communication system designers to provide basic services to applications, and to not anticipate the end use (or worse, only allow that one use). Of course, the Internet is the prime example of the right way to do this: the underlying TCP and IP protocols provide a basic service, and legislate nothing about what can travel above them. The intelligence of the network is at the edges, not in the connections. Other programmers used this freedom to create all sorts of unforeseen services: streaming video, the world-wide web, secure shells, etc. In contrast, the current generation of digital phone networks allow no innovation. If you had an idea for a great application for a mobile phone, you'd have to sell it to the phone company or forget about it.

It's easy to understand why this is true: money. More specifically, monetization. A corporation interested in building out a wireless service has to invest a staggering amount of money in physical towers, bandwidth rights, or rather famously satellites. According to the Wall Street Journal, the telecommunications industry has amassed more than $650 billion in debt. To raise that sort of capital, it needs to make a case that it will be able to recoup that investment. By controlling the entire vertical stack of applications, mobile service providers can promise a revenue stream, especially from high-margin add-on services.

Unfortunately, this top-down, centrally managed architecture will likely doom the effort. Only by allowing and embracing outside innovation will these networks flourish. But if vertically-integrated networks will fail, and

non-vertically-integrated networks won't get funded, how are we going to get our global, mobile, always-on world?

The best chance seems to be to encourage the creation of bottom-up networks, rather than monstrously expensive top-down, centrally managed networks. These bottom-up networks would emerge from the piecemeal action of many users, rather than a few telecom providers, and would be built on standards and open protocols. The best candidate for this in the wireless space is the emergence of the fluidly-named 802.11b, a technology for wireless local-area networking. The potential is that rather than being a local-area net, as its inventors intended, it might actually become the basis of a disconnected wide-area network, providing services to mobile users in any office, airport, or hotel.

The first necessary ingredient is local overcapacity. When local offices install 802.11b, they do so to meet their own needs. But the fact is that for almost all installations, there's excess capacity. This is simply due to the fact that wireless capacity generally comes in large chunks: a single base station generally provides access to a lot of bandwidth. Once this local overcapacity exists, it becomes possible — and economically convenient — to share that excess capacity. This is the same dynamic, applied to disk drives and wired bandwidth, that drove the success of Napster.

The second ingredient is to encourage reciprocity and cooperation. If there's local overcapacity in multiple locations, and users are mobile, it makes perfect sense to make an arrangement to share. Sharing this overcapacity makes economic sense, but technical needs must be met as well. For wireless networks, this means strong authentication, per-client security, and possibly a convenient micropayment or subscription basis.

Bottom-up networks, like kindergarteners, must play well with others. This means that new networks must at the very least stay out of the way of others; better yet is to interoperate cleanly. The alleged collision between Bluetooth and 802.11 is an instance of a stumble that may have delayed rollout of the technologies for a year, if it didn't kill

Bluetooth outright. In contrast, the introduction of 802.11a is carefully designed to avoid such conflicts, but it isn't interoperable with 802.11b. Will anyone carry two wireless cards when traveling?

This issue of sharing the spectrum is a profoundly difficult one. As posed by Larry Keeley of the Doblin Group, if we had to design the rules for spectrum sharing from scratch (say, on a new planet), we wouldn't know how to do it right. In this much more complicated world where there are already thousands of stakeholders, it's all the more difficult.

One innovative path to a solution is to move spectrum regulation away from a "Ten Commandments" model, where rules forbid certain activity, to a "Bill of Rights" model that gives devices certain rights. For example, rather than forbidding broadcasting at more than a certain power in a certain frequency, this model would allow a device to broadcast at any power in any wavelength as long as it didn't interfere with any other device. Could this work? Would the FAA ever agree to allow even microtransmitters to broadcast in airplane navigation frequencies? It seems unlikely, but it remains an intriguing and thought-provoking proposal.

The final ingredient to get right is, of course, standards. Interoperability, cooperation, and reciprocity all require the establishment of standards, or at least agreements on protocols. The problem comes in that establishing standards too early (as in the HDTV case) risks missing the necessary innovation and change. On the other hand, if you wait too long to standardize (as the United States may have done with wireless providers in the 1990s), you lose out on network effects, both large investors and consumers hold back until a winner emerges, and limited effort and investment is diluted over many competing systems. As Jared Spool points out, at some point you need to shoot the innovators. The fax machine was invented in 1843, but it wasn't until 1983, when a single standard way to send a fax to any machine was established, that they exploded in popularity. Moving from innovation to standardization at precisely the right time is a key factor of success.

Conclusion

The wireless world really is full of potential - potential for profit, for exciting innovation, and for the chance to really change the way we work and play. But it also labors under enormous burdens. Telecommunication companies have spent billions acquiring the rights to monopolies on certain characteristics of electromagnetic radiation, and billions more building unsightly towers. But the technological challenges of bringing new services to market in a way compelling enough to command high prices may escape them, as it escaped Iridium. If 3G fails, the telecom industry runs the risk of having tried and failed too often, and being unable to effectively raise capital for wireless communication.

There are many technical challenges ahead. To name just one, real progress in integrating wireless communications with human life will require significant advances in how we automate intermediaries and negotiation. Humans just can't be bothered to inspect every wireless coupon, advertisement, or marketing call. Of course, there are challenges just as difficult in signal processing, efficient chip design, and battery chemistry.

But the way ahead isn't blocked by these technical challenges; eventually, some smart people will solve those problems. Yet there are higher hurdles to leap. To quote cartoonist Walt Kelly, we have met the enemy, and he is us. We simply don't know what to do with mobile devices yet. We need to learn — indeed, we need to create — a new language of abstractions to describe our needs and priorities to the network. We need to learn which functions really can fit in the same molded-plastic box, and which can't. We need to set rules about when it's okay to take a call. We need to struggle with powerful, emotional issues surrounding privacy, security, and anonymity. We need to learn to stop driving while talking on the phone. Learning these lessons may require more of a change in us than a change in the devices. In the end, of all of the travels that we go on with these devices, the networked world may be the most foreign.

For more information, contact Geoff Cohen at geoff.cohen@us.cgeyc.com

About the Authors

Geoff Cohen is a Senior Consultant at the Cap Gemini Ernst & Young Center for Business Innovation in Cambridge, Massachusetts, where he researches trends in emerging technology and innovation. He has also worked for the Congressional Budget Office, Data General, and IBM. He holds a Ph.D. in Computer Science from Duke University and a B.A. from the Woodrow Wilson School of Public and International Affairs at Princeton University.

Mukul Kanabar is a consultant at the Center for Business Innovation, where he works on the Valuing Intangibles and Networked Commerce initiatives. His work on the intangible valuation initiative is focused on understanding and quantifying the impact of non-financial measures (brand, management quality, customer satisfaction, etc.) on an organization. Further, he is helping to develop tools that organizations can use to measure, manage and report these non-financials. Mukul also conducts research on several technology related fronts. In particular, he has been researching the future of wireless devices, data and networks and the use of real options in technology investment. Mukul holds a B.A. in Economics from Haverford College.