



Alec Saunders

Voice 2.0: A Manifesto for the Future

We're witnessing the beginnings of a titanic clash between the internet and the telecommunications industry. My hope is that clash will be the, albeit painful, evolution of Voice into a full blown internet application — the birth of Voice 2.0. Voice 2.0 is the next step from where we are today. In today's world, VoIP "carriers" like Vonage, Packet8, and the cable offerings, are migrations of the legacy PSTN onto a VoIP foundation. Voice 2.0 — true VoIP — is the marriage of IP Telephony to the Web.

It's already begun. The arrival (and mass adoption) of technologies like Skype, Peerio, and PhoneGnome are one indicator. Another is the accelerating loss of landline business amongst incumbent carriers. In the first quarter of this year, North American landline attrition doubled. As I write this today, landline cancellations have reached 10,000 per day, as customers opt for cable, mobile, or VoIP solutions over services from traditional providers.

What will that world look like? Who will be the winners, the losers, the moneymakers? What will the consumer experience be? Follow along with me, and let's have a closer look. This essay is part fiction, and part reality. It's a whole lot of what I would like to see in the communications platform of the future, which I have dubbed Voice 2.0.

Talk is the baseline

In a typically Scandinavian understatement, Skype founder Niklas Zennstrom's rationale for why Skype is so successful was this: "People like to talk". People do like to talk. As communications services have become ever cheaper, the explosion of usage has been remarkable. For instance, according to the FCC, between 1994 and 2001, average minutes of wireless usage per month climbed from 140 minutes per month per user, to 427 minutes per month per user. Over the same period of time, annual US land line usage grew from 2.8 trillion to 4.8 trillion minutes. Prices fell, usage skyrocketed.

The merger of talk with the web is the foundation of Voice 2.0. When Skype launched, and the price of minutes dropped to zero, social barriers to calling strangers disappeared, driving voice usage higher again. The merger of talk and the web is leading to web based conferencing, push to talk, application sharing, voice enabled e-commerce, and a multitude of other applications, all of which are driving voice usage higher. In the process this merger is redefining the staples of business — customer service, sales, and marketing — and impacting all of our lives as we move from the standard work day to 24/7 availability.

Talk is the baseline, but that baseline will be combined with text / IM messaging, and video. Today's networks can support the technologies. The evolution to full blown, multimedia, real time communications is just a matter of time. Some products, like the Nokia N90 cellular telephone, are already providing this capability. Nokia's E Series telephones will also have built in SIP clients, facilitating a seamlessly mobile VoIP world.

As speculative fiction writer William Gibson said, "The future is already here, it's just unevenly distributed." It begins with talk.

The meter is off

Voice will be free, as the Skypers contend, and the Stupid Network model implies. Short term, all you can eat models, like Vonage, will exist, but long term it's clear that the metered model is dead. The point-to-point technology called VoIP neither requires, nor facilitates the metering of traffic. Metered access to mediated access networks, like the PSTN, will continue only so long as customers require access to those networks to talk.

Currently, the only widespread metered model in VoIP is metered access from the IP network to the PSTN. But how long before the majority of customers are on the IP network, and the model reverses? When will we see PSTN customers pay a premium to contact their friends on VoIP networks?

In the Voice 2.0 world, there will be three billable entities: connectivity, directory, and applications. Connectivity and directory will be low margin, commodity businesses. Customers will pay for access to the network, and perhaps to be listed in a directory. Applications will be the value creators. The meter is off.

Applications as the value creators

If talk is the baseline, then applications will be where value is created. Applications in the Voice 2.0 world will come in three flavours:

- Voice applications - these are the traditional voice applications we already know and love. IVR, unified messaging, conferencing, and perhaps others. Sophisticated new tools, such as VoiceXML, will provide ways to create richer and more powerful voice applications, and drive further value.
- Voice enabled IT applications - these are intersection of today's business process automation tools with voice. Sales force automation, CRM, accounting, email, payroll, etc. Every application we use today will become communications enabled with voice. Next generation softphones will evolve from simple replacement systems for road warriors into full blown platform components for knowledge worker applications
- The voice web - the mash-up of voice and the internet will result in a whole new class of applications, not seen yet. VoiceXML is the first step in this direction. Tools like PHPVoice take this a step further, allowing for sophisticated voice applications to be built using the same kinds of scripting tools that the web is built from. However, the real pot of gold is the combination of web services and voice. Examples include: spoken word real estate descriptions from the MLS coupled with mapping, voice enabled matchmaker services, customer service coupled with inventory / ordering / availability. The mix of text, web, voice, and programmatic access to data is a heady brew.

The construction and sale of these applications will be a market bigger than the web itself. Applications are where the value is created in Voice 2.0.

Building blocks: presence, directory, XML

Much of the talk in VoIP is about new kinds of media — video, wideband voice, IM. Media may be the star of the new network, but the workhorses of Voice 2.0 applications are signaling and control components. Presence, to determine availability; directory, to determine addressing and routing; and XML web services for call control, and integration with computing assets. These are the true value creation components in the architecture.

Presence will drive a fundamental change in the way that communications networks are used today. Today, callers have no way of knowing whether the party being called is available, or busy, or would consider the call an intrusion. With presence, the availability of a called party is known by the calling party before making the call. This seemingly simple idea will increase the immediacy and value of calls in all kinds of applications. It will do away with calls that begin with "Is this a good time", and reduce the volume of voice mail created when parties can't connect. The range of new applications for presence is huge also. For instance, imagine an ad-hoc collaboration application where you are able to know, before initiating the call, whether all parties can attend the call, right now.

Directories have existed since the advent of voice networks. However, in the Voice 2.0 world, individuals own their own directory listings. What you wish to list in your directory listing, including the fundamentals of name, address, and contact point(s), is your business. It's your identity, and you get to manage it — not the carrier. Directories can be extended to include the idea of persona's (work, home, leisure), interests, and a myriad of other kinds of personal information. Directories also become repositories for subscriber preferences, credentials, social networking details and potentially even financial information for voice enabled transactions. In the voice 2.0 world, the directory is an opt-in enabler for applications, commerce, and identity.

And last, but probably not least, are **XML based web services APIs** for accessing presence, call control, and directory. Scriptable building blocks have been part of IT for over a decade. One of the primary successes of Web 2.0 has been the extension of scripting to the internet, with XML based APIs used to create mash-ups between applications. Scriptable communications building blocks are the next step. In the voice 2.0 world any application, within the bounds of permissions set by the subscriber, can access presence; initiate, accept, and redirect calls; and query directories. Experiments today, like Free World Dial-Up's exposure of CPL and CCXML, and Skype's partnerships with VoiceXML providers, are pointing the way.

New value networks

The business implications of creating an open, programmable web services architecture for voice are profound. When service provider architectures are open, the envelope — the monthly envelope that includes the integrated bill for all the services you buy from your service provider — disappears. Where does a service provider's network begin and end in this world? How do you monetize the component building blocks? The applications? Should you limit access to the building blocks or make them entirely open?

Voice 2.0: A Manifesto for the Future

Will subscribers pay for the platform building blocks? No. Platform components, in and of themselves, are not interesting to consumers. Applications are. Revenue from the applications must be shared with building block owners in order to facilitate growth of the platform. The business model is built on settling transactions at the interfaces between applications and building block owners, rather than the interfaces between networks.

Fundamentally, this turns the service provider value network on it's head. In today's world, the network operator aggregates services from a number of vendors, and then delivers them to the customer. Tomorrow, the customer will buy the services they want from whomever they want, and the service provider will deliver a portion of that revenue to the owner of the platform component. Voice will be monetized through the long tail of high value applications targeted at specific communities of users.

In the Voice 2.0 world, applications are ascendant, and platform and network components are buried in the applications.

The shape of things to come

Voice 2.0 is a user-centric view of the world. In Voice 2.0, "it's all about me" — my applications, my identity, my availability. Voice 2.0 is all about developers too — the companies that exploit the platform assets of identity, presence, and call control. It's not about the network anymore.

All of the technical underpinnings I've described so far exist today. One element still missing is a common, standardized, presentation layer. Standards exist for this layer — VoiceXML, SIP, SALT etc all read on presentation. However, at this point there is no ubiquitous equivalent of the HTML browser. The closest yet are Skype (completely proprietary) and Gizmo Project (well below ubiquity, but very complete standards implementation). It's most likely that one of the VoIM players, (Microsoft, Yahoo, Google, AOL) will drive this.

And what of the PSTN stalwarts, the incumbents? For some, their time is done. Others have made consistent and intelligent investments in next generation technologies. The challenge for those will be The Innovator's Dilemma — how to transition into this new world, while maintaining profitability, and retaining shareholder loyalty. It will take a deft hand, indeed.

