

The Fifteenth Annual Interactive Audio Conference

PROJECT BAR-B-Q 2010



Group Report: Wherever You Go, There You Are: Audio that Understands Context and Mobility

Participants: *A.K.A. "Nomadic Hoarders"*

Peter Kirn, Create Digital Media

Moshe Sheier, CEVA

Devon Worrell, Intel

David Roach, Optimal Sound

Ron Kuper, Sonos

Dawn Leonetti, Dolby

Facilitator: Jim Rippie, Invisible Industries

The problem:

Diverse media sources, devices, and contexts fragment the experience of using sound.

A proliferation of devices and media sources has made accessing your music complicated and cumbersome.

Additionally, audio routing today through multiple devices incurs multiple signal processors that may not be aware of each other and may be incompatible.

Content is spread across a variety of sources, streamed and local. We use a variety of devices, each of which mixes music listening and voice communication functions. There's no consolidated access to all of those sources, and no information about what you're doing or where you are doing it. As you move from place to place, you want your audio to be associated with you, not only content sources and devices.

Various components attempt to solve pieces of this problem. UPnP, DLNA, and zeroconf (Bonjour) address device discovery. Aggregators like lockers, RadioTime, Rovi, and Muse consolidate some media sources. Services like Last.fm track user listening data across players and devices. But these services don't talk to each other, and the sum total of the services do not create or share enough information to allow seamless playback across systems and devices. This content will also need to be socialized and shareable with others in our social networks or living environments.

The solution:

Use the cloud to make audio devices aware of where you are and what you're doing.

We believe the remedy to these problems centers around an online, connected service that catalogs a user's devices, content and listening environments.

We propose a cloud-based service that provides:

- A continuously-updated inventory of the user's media and communications, including sources and content identifiers.
- A catalog of a user's devices and capabilities, including the presence of DSP algorithms in the signal chain and whether they can be configured.

- Contextual information with each activity, including where, when, and with what devices it occurs. If desired, this can also layer richer information about the associated environment.
- Profiles for the user's contacts such as "in the car," "in the living room," "at a friend's house," "at work," or "on a bike." Like assembling a playlist, the user could make use of open-ended, more meaningful profiles like "on vacation in Iceland."

Using this service, the user can move seamlessly from context to context, by automatically performing one of two basic operations:

- Handing off the stream or content from one device to another, so that a media play can continue without significant interruption. The client will load media locally from a device if available, or will choose a different, connected source if it is not.
- Re-directing or forwarding an existing stream to the other device, much like the functionality of the Slingbox for video, enabling media and conversations to continue without restarting the stream.

In addition to building the cloud system, client apps would need to be developed to access this service.

Clients with DSP capabilities can use the DSP catalog on the service to prevent conflicts between processing algorithms.

See previous Project BBQ work for more thoughts on this issue; most recently, "**I Hear The Future: The Binaural Headset as Audio Contact Lens and Our Inevitable Mixed-in Lifestyle of Personal Audio Networks**" (2007), "**Smart Ambient Sound Sensor**" (2008), "**Here, There, and Everywhere**" (2009), and "**Mobile Infrastructure**" (2009).

Action items from workgroup

Peter Kirn, assisted by Jim Rippie – complete report for publication

Moshe Sheier - investigate sensing methods for detecting how a user moves between devices, environments, and contexts. Initial report/early results.

Peter Kirn - experiment in sensing music playback on Android.

Ron Kuper - present an implementation road map for the aggregator.

Steve Tellman - catalog a range of devices and capabilities

Devon Worrell - lobby Microsoft, Apple, Google; catalog use cases for report.

Dawn Leonetti - investigate social network capabilities and how they integrate (high level functional view) with use cases

Addendum:

Sensing methods for detecting how a user moves between devices, environments, and contexts

Moshe Sheier, (additional device hardware capabilities) Peter Kirn

Two main methods could be used for observing a user movement between devices, environments, and contexts – manual (no sensing involved) and automatic.

Manually ("act"), a user would register for our aggregator service (described elsewhere in this report), using the device he would like to listen through. Only one device could be registered at a time. This way, once registered thorough the user's home, office, or mobile device, the last played music track would continue on the newly registered device. While this is a straightforward approach, it adds the hassle of registering every small device/location change we make, and we would also like to consider a more automatic way of moving around.

An example for manual location aware music system (using a web service) is described in the following paper: <http://www.drhu.eu/publications/2010-ICCE-LsM->

[ANewLocationAndEmotionAwareWebbasedInteractiveMusicSystem.pdf](#)

Automatically (“react”), several methods could be used to sense a user location:

RFID – in a per-configured environment (home/office) a user could be equipped with a RFID tag that would get sensed at each room (using appropriate RFID readers). Music would be played to the room the user is in, based on the audio equipment at that room (recall that we assume all audio equipment is “connected”).

An example is shown in this paper - <http://ieeexplore.ieee.org/application/enterprise/entconfirmation.jsp?arnumber=5319207>

By extension, advances in NFC (Near Field Communication) wireless sensors, expected to be widely available in Android mobile devices (as publicly demonstrated for Android's “Gingerbread” release) and likely other mobile platforms (iOS, etc.) would make the potential for automatic sensing ubiquitous. Hand-off could be designed into an intuitive gesture, like swiping a cell phone NFC device across an NFC-equipped car audio system in order to pass off a conversation to hands-free speaker operation, or music playback from headphones to the car. The importance of having a commonly-understood protocol for communicating sound use “intents” irrespective of platform would therefore provide the users' ability to traverse various sound contexts.

GPS – as our SmartPhone is equipped with a GPS receiver, it is an ideal method to determine our location while on the move. Whenever our aggregator system get's notified that we are out of the home/office location, music should be played either through the SmartPhone itself, or through a Bluetooth add-on (e.g. a car audio system) paired to our SmartPhone.

In addition to sensing a user location/device being used, the environment the user is in should also be sensed, in terms of context (quiet office, noisy party), other people in the room, and their music preferences.

Some work is being presented here – http://www.pervasive2006.org/ap/pervasive2006_adjunct_1E.pdf

Aside from obvious (and precise) GPS, onboard hardware sensors available in mobile devices (and increasingly available not only via native hardware APIs but next-generation mobile browsers, as well) can provide other clues to context:

- Microphone, for ambient sound levels
- Motion sensors, showing whether the user is stationary or moving
- Proximity to a face, for voice interaction
- Camera, sensing movement, faces, and other contextual information

By combining automatic sensing and manual registration (where necessary), our system would become aware of the current device a user plans to use for audio playback and the environment it is in, for our nomadic audio experience to be utilized.

Illustrations and Use Scenarios

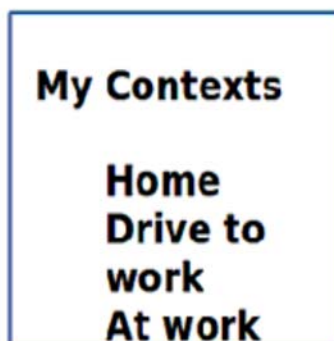
Ron Kuper

Smart(er) cloud, dumb(er) client: In one implementation case, intelligence is built into cloud, for a centralized catalog of user devices, content, and contexts.

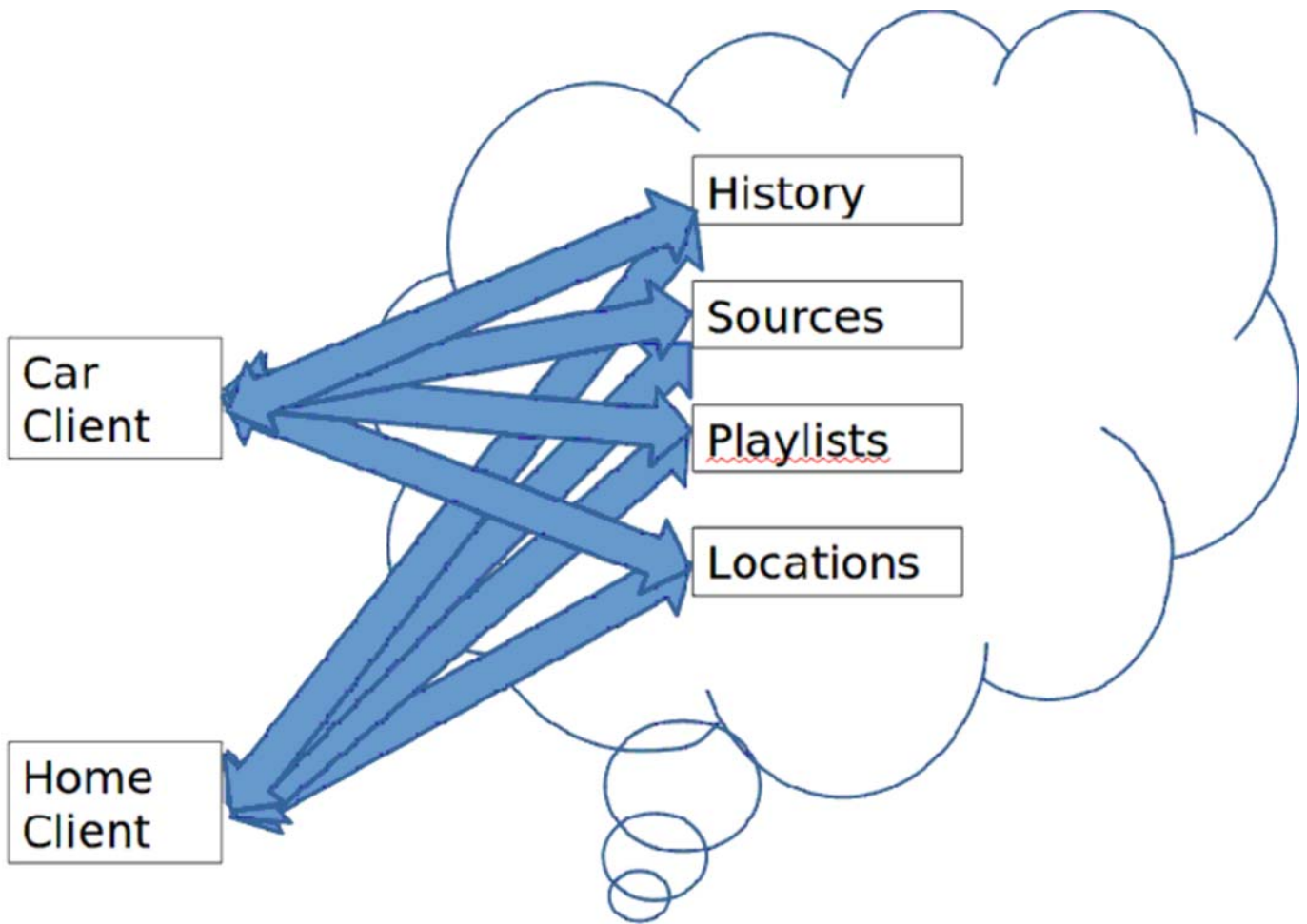
In usage, a single activity – like going to work – raises a number of queries for that cloud-based catalog:

- What's the content? (“What am I listening to?”)
- What's the device? (Does a switch in context require a change in device?)
- How can content move to the new device uninterrupted? (Are the two devices in sync? Can they be synced? Can playback be handed off to the new device? If not, can a stream be forwarded to the new device?)

Once those issues are resolved, a resulting action is taken – like setting up, then forwarding a stream from the old device to the new.



Dumb(er) cloud, smart(er) client: In our second implementation case, the cloud is still an essential ingredient, but greater intelligence is built into the client. (Indeed, the expanding array of sensors described elsewhere in this report suggest one reason such a scenario might come into play. The local device may be able to sense more information and immediate context than the cloud could collect or anticipate.)



Social interaction

Regardless of the most appropriate breakdown of client and server, keeping data in the cloud allows the user to collect and selectively share data about sound and context. In addition to common use cases that already exist – like publishing a currently-playing track to an instant messaging status field or discovering common musical tastes – this opens up new possibilities. A friend could see that someone is currently on a conference call and unavailable. Someone caught navigating traffic in an unfamiliar city could indicate the desire to remain uninterrupted. Where users do want to share information about their devices, content, and context, they can, to those with whom they wish to share.

Social - Lots of happy clouds

