WebRTC and speech recognition services with Adhearsion

Luca Pradovera
FOSDEM 2017
WHO AM I?

• Luca Pradovera
• New Principal/Lead at Mojo Lingo LLC
• Adhearsion contributor
• Played with phones since I was 8
DEMO FIRST! (SOMEONE CALL JAMES BODY)
WHAT WAS THAT?

The demo might not actually contain WebRTC.
Consult your physician before attempting to configure WebRTC on a local machine.
No keyboards have been harmed during the preparation of this demo. Honest.
MOVING PARTS (ALL OPEN SOURCE)

- FreeSWITCH and mod_verto
- Adhearsion
- PocketSphinx
- Flite
- Rasa NLU
- …and a bunch of others
WHAT IS FREESWITCH?

• SIP-based PBX
• Tons of features
• Very modular
• Very good WebRTC support through mod_verto
• Also check out Asterisk
THE BOT’S EAR AND VOICE

• PocketSphinx provides ASR
• Could be tuned for better results
• Flite provides TTS
• Of course you could use others
THE BOT’S BRAIN

• Rasa NLU is a very interesting NLP and ML library
• It replicates services such as Wit.ai, LUIS and Api.ai
• Compatible with many formats and learning models
• We are using the restaurant demo
• https://github.com/golastmile/rasa_nlu
WHAT DID I LEARN BUILDING THE APP?

• We need a better way to set up FreeSWITCH or Asterisk for WebRTC development
• PocketSphinx is not as bad as the reputation it has (YMMV)
• There is value in running your own “brain”
• Adhearsion removes a lot of complexity
WHY USE ADHEARSION?
WHAT IS ADHEARSION?

• Ruby voice application framework
• Provides 3PCC logic to telephony engines
• Connects to FreeSWITCH using Rayo, to Asterisk using AMI
• Version 2 is stable, version 3 is at rc1
• Backed by Adhearsion Foundation
WHAT IS NEW IN ADHEARSION 3?

• FreeSWITCH support is Rayo only
• Asterisk 11+ required
• Streamlined internals
• Built in HTTP server
• Native i18n support
WHAT DOES ADHEARSION PROVIDE?

- Plugin architecture
- Voicemail, pseudo-TTS, call queuing plugins
- Platform-specific functionality plugins
- Unified logging
- Clustering via Rayo
- Better deployments using Ruby standards
HOW DOES ADHEARSION WORK?

• Represents phone calls as actors
• Passes messages and events between the engine and the actors
• Each call runs its handling logic in the actor thread

```ruby
class MyController < Adhearsion::CallController
  def run
    answer
    resp = ask "How many woodchucks?", limit: 1
    say "You said #{resp}. That's obviously wrong!"
  end
end
```
GENERAL APPLICATION STRUCTURE

• Controllers group up features
• Routing controls which controller gets a call
• An event handler catches server messages
• Based on Celluloid, operation is generally async and event-based
• DSLs for all common operations (playback, recording, menus)
RAYO PROTOCOL

- XMPP based 3PCC protocol
- Encapsulates voice app primitives
- First-class citizen in FS through mod_rayo
- Calls, speech and TTS, mixing, media
- As a side effect, every Adhearsion node has an XMPP address

http://rayo.org/
ADHEARSION ON ASTERISK

- No Rayo support
- Connects via AMI
- Has native command support
- Slightly easier to get started
WHAT CAN I DO?

- Calls, conferences
- Media with I18N
- Drive GRXML/SSML based ASR/TTS
- Complex IVRs
- API calls
- Database access
- Built in HTTP server
- Not limited to the dialplan

Everything but the...
HOW IS IT DEPLOYED?

• Any Ruby flavor
• Usually 1-1 with FreeSWITCH
• 12-factor compatible Ruby process
• Easier to scale, provided you have a load balancer
CODE COMPARISON: XML DIALPLAN

• Simple to build
• Nothing to manage
• Difficult to integrate
ADHEARSION CONTROLLER

• Code reuse
• Ruby Gem ecosystem
• Complete language
GIVE US SOME EXAMPLES!
CASE STUDY: RINGRx

• The only HIPAA-compliant phone system
• A cloud PBX and an On-Call service
• Features handled by Adhearsion:
  • Conditional routing
  • Voicemail recording and moving
  • Custom message recording and custom IVR
  • Reminder calls
  • …pretty much everything else.
CASE STUDY: LiveConnect

- Surgical procedure broadcast system
- SIP-based because of hardware
- One SIP broadcaster, N WebRTC (mod_verto) or SIP clients
- Adhearsion used for:
  - Managing security and access
  - Conference room participants
  - HTTP API to control flow switching
  - Recording handling
CASE STUDY: POWER HOME REMODELING

• Home renovation company
• 400 Call Center operators
• Outbound for sales and appointments
• Inbound for field agent and installation support
• Every business is a communications business
MORE EXAMPLES?

• Major publishing company phone system for handling delivery accounts, complaints, and services
• At least one MVNO (guess which one)
• Cultural mediator network with online translation
THANK YOU!

My name is: **Luca Pradovera**
I am a Voice Application Developer
at **Mojo Lingo**.

Web: [https://mojolingo.com](https://mojolingo.com)
Twitter: [@lucaprado](https://twitter.com/lucaprado)
GitHub: **polysics**