Speech transcription and intelligent CRM event triggering

Desjardin RFP
September 2019

This document is an overview of the solution design for the speech transcription component of the larger CRM proposal. This functionality is used to address some of the needs outlined in Use cases 1 & 4 of Annex 2.

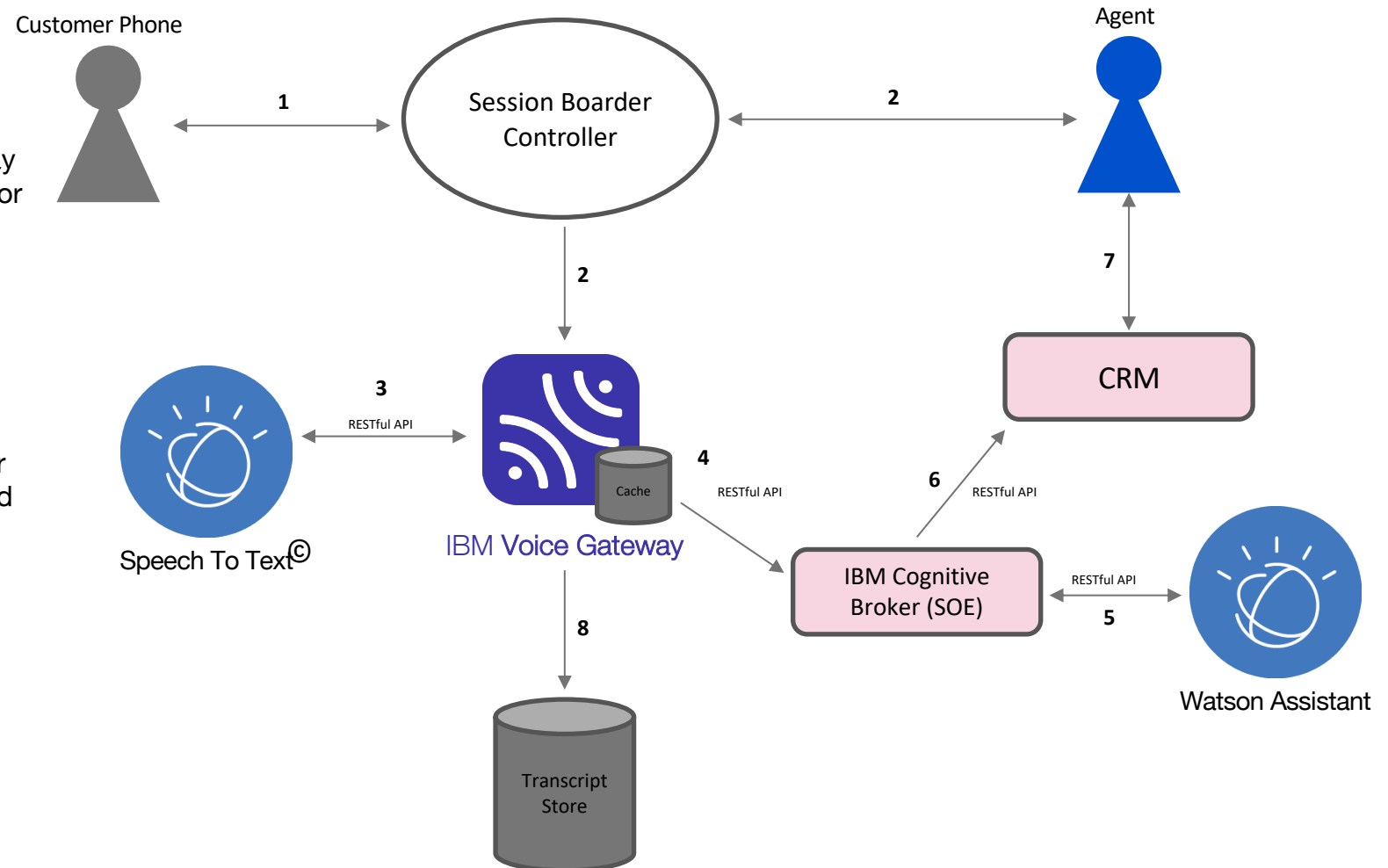Two solutions have been provided (1) Real time processing via telephony (2) Post Call recording processing.

The document lays out the following :

▪ High Level Solution design

▪ Assumptions

▪ High Level Implementation Tasks

▪ Effort Estimate (in hours)

▪ Rough Cost Estimate

Option 1: Real Time

Audio streamed in real-time for transcription, CRM updates and agent assistance

1. Customer initiates call
2. Audio is split to Desjardin Agent and to IBM Voice gateway
3. Voice gateway streams audio to Watson Speech to text for transcription based on language identified by telephone channel, transcribed sentences are returned to gateway and cached during call
4. Recording is submitted to Watson Text to Speech to transcribe voice for both agent and customer, call meta data identifies call language to be transcribed. Full transcript is returned to orchestrator
5. Transcribed sentence is submitted to Watson assistant for text classification and to identify update event triggers and returned to SOE
6. Any updates that require fulfillment to CRM will be posted via RESTful interface
7. Updates are pulled to agent desktop application
8. Once call has been terminated the full call transcription is submitted to transcript storage from cache

Customer Phone

Agent

Session Boarder Controller

1

2

2

7

CRM

3

RESTful API

Speech To Text©

IBM Voice Gateway

Cache

4

RESTful API

6

RESTful API

IBM Cognitive Broker (SOE)

RESTful API

5

Watson Assistant

8

Transcript Store

* Solution can support various STT platforms such as Google, Nuance, Amazon etc.

Option 1: Real Time – Effort Estimate

Assumptions

1.Assumes no more than 20 Intents training in both English and French
2.Prerequisite hardware and software has already been procured and setup outside the IBM responsibilities listed below
3.The call center will not have over 500 concurrent calls at any one time
4.SBC configuration to trigger IBM Voice Gateway is reliant on Desjardin
5.Solution will be deployed on Desjardin premise using the IBM Cloud Pack for Data (ICP4D)
6.Will leverage the existing IBM Cognitive Broker asset for integration with CRM solution
7.Details from phone call will be transcribed and leads will be created on the CRM during the call.
8.Solution will be deployed to Dev, Test and Prod environments with
9.Solution will support High Availability
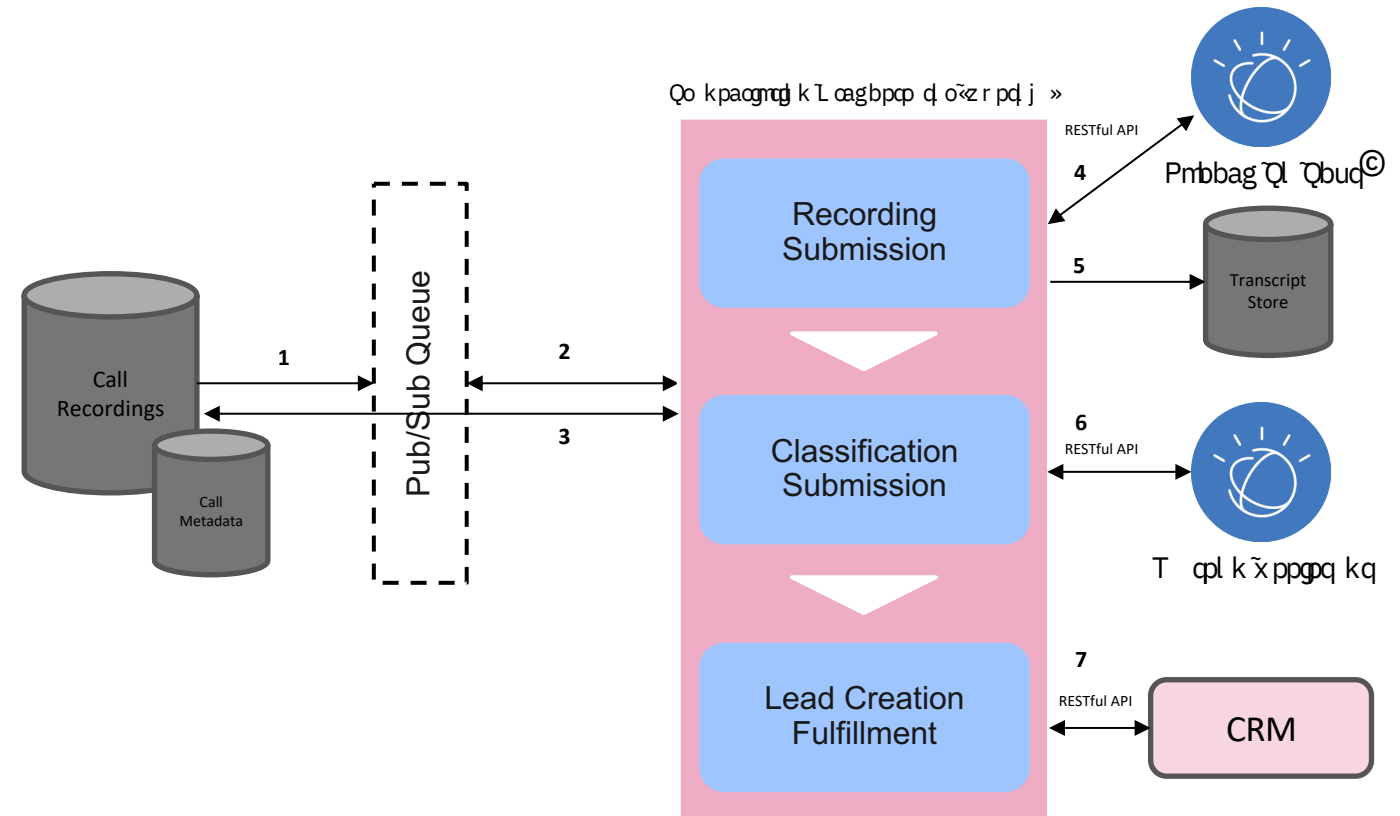Natural Language Classification will require continuous improvement with production data following implementation

Tasks & Services Effort

Task	Description	Services Effort (Hours)
ICP4D Setup for Dev , Test & Prod – HA	ICP4D set up includes the creation of 3 environments in HA using IBM recommended topology, Integration to LDAP compliant user registry, Watson Assistant and STT Setup and functional testing	240
Voice Gateway Setup	Deployment of Voice Gateway solution to Dev, Test, Prod ICP environment, integration to STT and IBM Cognitive Broker, Mapping to test phone number, Integration to Data store and integration testing	530
IBM Cognitive Broker Setup	Deployment of IBM Cognitive Broker to Dev, Test and Prod clusters, integration to CRM Solution	240
Watson Assistant Training	Development of Natural Language Classifier based on historical call data, development of CRM update fulfillment workflows in both English and French	240
STT Custom Training	Training of custom grammar and terminology in English and French, custom acoustic model	200
Total Hours		1450
Total Cost Estimate +- 20%		$375,000

Option 2: Post Call

Full audio transcription occurring post call to identify updates to CRM

1. Call recordings are saved to databased once a call is terminated and message of new recording is published to Pub/Sub Queue
2. Transcription orchestrator is subscribed to queue and utilizes the key provided in the message to pull the recording
3. Orchestrator pulls call recording and call meta
4. Recording is submitted to Watson Text to Speech to transcribe voice for both agent and customer, call meta data identifies call language to be transcribed. Full transcript is returned to orchestrator
5. Transcript is stored in database with associated call meta data
6. Customers transcript sentences are separated and submitted to Watson assistant individually for text classification and to identify update event triggers and returned to Orchestrator
7. Any updates that require fulfillment to CRM will be posted via RESTful interface

Call Recordings

Call Metadata

Pub/Sub Queue

1

2

3

Recording Submission

Classification Submission

Lead Creation Fulfillment

RESTful API

4

5

Transcript Store

RESTful API

6

RESTful API

7

CRM

* Solution can support various STT platforms such as Google, Nuance, Amazon etc.

Option 2: Post Call – Effort Estimate

Assumptions

1. Assumes no more than 50 Intents training in both English and French
2. Prerequisite hardware and software has already been procured and setup outside the IBM responsibilities listed below
3. Call audio is stored with call meta data in an existing storage and can be published to a Pub/Sub Queue for consumption in the transcription and fulfillment pipeline. Recordings will be in either mpg3, opus, mpeg, wav or flac file formats only.
4. Solution will be deployed on Desjardin premise using the IBM Cloud Pack for Data (ICP4D)
5. Solution will require the creation of a custom orchestration engine for the transcription and fulfillment pipeline
6. Solution will be deployed to Dev, Test and Prod environments with
7. Solution will support High Availability
8. Natural Language Classification will require continuous improvement with production data following implementation

Tasks & Services Effort

Task	Description	Services Effort (Hours)
ICP4D Setup for Dev , Test & Prod – HA	ICP4D set up includes the creation of 3 environments in HA using IBM recommended topology, Integration to LDAP compliant user registry, Watson Assistant and STT Setup and functional testing	240
Pub/Sub Queue Setup	Deployment of queue to Dev, Test, Prod and functional testing	240
Development of Transcription Orchestrator	Deployment of custom orchestration application to submit recordings for transcription, natural language classification and to fulfill CRM updates. Deployment will include integration testing in dev, test and prod environments	240
Watson Assistant Training	Development of Natural Language Classifier based on historical call data, development of CRM update fulfillment workflows in both English and French	240
STT Custom Training	Training of custom grammar and terminology in English and French, custom acoustic model training	200
Total Hours		1160
Total Cost Estimate +- 20%		$300,000