

---

## Title: How audio is processed with a controller weight: 50

Typically, you use the local speaker and microphone of the audio device to process audio. However, you might have a smart speaker in your environment that you would like to use to perform external processing or you might have additional device controls, for example, for volume or display.

Watson Assistant Solutions provides you with the option to use your own smart speaker and microphone with the audio client. The smart speaker acts as a device controller for the audio client. You use a command socket interface to send commands between the device controller and the audio client and an audio socket interface to send audio.

In this scenario, the device controller listens for a wake up command, instructs the audio client to start listening on the audio socket port, and passes the audio input from the end-user to the audio client.

The audio client passes an IAM token when it first connects to audio gateway to authenticate. For more information about setting up IAM authentication for an audio client, see the [Authenticating audio devices](#) topic.

The flow for processing audio input in Watson Assistant Solutions when a device controller is used is shown in figure 1.

Figure 1 - audio flow with a device controller  Flow with a device controller

The steps in this flow are as follows:

1. When it turns on, the audio client sends an IAM token to the audio gateway and the gateway authenticates the client.
2. The device controller processes a wake-up command.
3. The device controller opens its microphone and listens.
4. The device controller receives audio input from a user utterance.
5. The device controller sends a **RAS** message to the audio client to instruct it to start listening.
6. The device controller sends audio to the audio client on the audio socket interface.
7. The audio client reads audio from the audio socket interface.
8. The audio client sends an **audio start** message to the audio gateway.
9. For each chunk of binary audio data in the utterance, the audio client sends the audio that it captures in an **audio data** message. The audio gateway converts the audio input to text in real-time using a speech-to-text service.
10. The audio gateway sends **speech-to-text transcript** messages to the audio client.
11. The audio client sends an **audio end** message to the gateway.
12. The audio client sends a **micOff** to the controller to turn off its mic.
13. The audio client sends a **micWakeupNotAllowed** to the controller to instruct it not to respond to additional wake-up commands for now.
14. The audio gateway forwards the text input to the routing core.
15. The routing core sends the utterance to the conversation component for evaluation and processing. The skill with the highest confidence processes the request and sends a

response. The conversation component forwards the converse response from the skill to the audio gateway. For details of these steps, see [How routing works](#).

16. If some or all of the response will be delivered to the user as audio, the audio gateway converts the response from the routing core to audio using a text-to-speech service. The gateway sends each chunk of binary audio data to the client in real-time.
17. The audio client sends audio to the device controller on the audio socket interface.
18. The audio client sends a `micWakeupAllowed` to the controller to instruct it to allow interruptions that made using the wake-up command.
19. If some or all of the response will be delivered as text, the audio gateway sends the text reply to the device.
20. The audio client checks if the audio gateway set the prompt parameter to `true` in the audio end message. If `true`, the skill expects a response from the end-user. The audio client sends a `micOpen` to the device controller. Otherwise, the device controller waits for a new wake-up command.
21. The device controller plays the audio response or displays the text response, or both.

### What to do next?

Learn about [the audio streaming interface](#).