语音交互服务

API 参考

文档版本 01

发布日期 2025-07-26

华为云计算技术有限公司

版权所有 © 华为云计算技术有限公司 2025。 保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明

HUAWEI和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为云计算技术有限公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为云计算技术有限公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为云计算技术有限公司

地址： 贵州省贵安新区黔中大道交兴功路华为云数据中心 邮编：550029

网址： https://www.huaweicloud.com/

目 录

1 使用前必读.....1
1.1 概述.....1
1.2 调用说明.....2
1.3 终端节点.....2
1.4 基本概念.....4
2 API 概览.....6
3 如何调用 REST API.....8
3.1 构造请求.....8
3.2 认证鉴权.....10
3.3 返回结果.....13
4 如何调用 WebSocket API.....14
5 实时语音识别接口.....16
5.1 接口说明.....16
5.2 Websocket 握手请求.....16
5.2.1 流式一句话.....16
5.2.2 实时语音识别连续模式.....20
5.2.3 实时语音识别单句模式.....24
5.3 实时语音识别请求.....28
5.3.1 实时语音识别工作流程.....28
5.3.2 开始识别.....29
5.3.3 发送音频数据.....32
5.3.4 结束识别.....32
5.4 实时语音识别响应.....33
5.4.1 开始识别请求响应.....33
5.4.2 事件响应.....34
5.4.3 识别结果响应.....35
5.4.4 错误响应.....38
5.4.5 严重错误响应.....39
5.4.6 结束识别请求响应.....40
6 一句话识别.....41
6.1 http 接口.....41

版权所有 © 华为云计算技术有限公司

6.2 websocket 接口 48
7 录音文件识别接口 59
7.1 提交录音文件识别任务 59
7.2 获取录音文件识别结果 68
8 录音文件识别极速版接口 73
9 语音合成接口 80
9.1 语音合成 80
9.2 SSML 标记语言介绍 89
10 实时语音合成接口 94
10.1 Websocket 握手请求 94
10.2 实时语音合成请求 98
10.2.1 开始语音合成请求 98
10.3 实时语音合成响应 102
10.3.1 开始合成响应 102
10.3.2 语音合成结果响应 103
10.3.2.1 音频流数据 103
10.3.2.2 时间戳数据 103
10.3.3 语音合成结束响应 105
10.3.4 语音合成错误响应 106
10.3.5 严重错误响应 106
11 热词管理接口 108
11.1 创建热词表 108
11.2 更新热词表 111
11.3 查询热词表信息 114
11.4 删除热词表 116
11.5 查询热词表列表 118
12 附录 121
12.1 示例音频 121
12.2 获取项目 ID 121
12.3 获取账号 ID 123
12.4 获取 AK/SK 123
12.5 公共请求参数 124
12.6 公共响应参数 125
12.7 状态码 125
12.8 错误码 127
12.9 配置 OBS 访问权限 133

1 使用前必读

1.1 概述

欢迎使用语音交互服务（Speech Interaction Service ，简称SIS ）。

语音交互服务（Speech Interaction Service，简称SIS ）是一种人机交互方式，用户通过实时访问和调用API获取语音交互结果。支持用户通过语音识别功能，将口述音频、普通话或者带有一定方言的语音文件识别成可编辑的文本，同时也支持通过语音合成功能将文本转换成逼真的语音等提升用户体验。适用场景如语音客服质检、会议记录、语音短消息、有声读物、电话回访等。

语音交互服务所提供的API为自研API。

表 1-1 实时语音识别接口说明

接口类型	说明
实时语音识别接口	华为云提供的Websocket接口，主要用于实时语音识别。音频分片传输，服务器端可以返回中间临时转写结果，在最后返回最终转写结果。

表 1-2 一句话识别接口说明

接口类型	说明
一句话识别	一句话识别接口，用于短语音的同步识别。一次性上传整个音频，响应中即返回识别结果。

表 1-3 录音文件识别接口说明

接口类型	说明
录音文件识别	录音文件识别接口，用于转写不超过5小时的音频。由于录音文件转写需要较长的时间，因此转写是异步的。

版权所有 © 华为云计算技术有限公司

表 1-4 语音合成接口说明

接口类型	说明
语音合成/实时语音合成	语音合成/实时语音合成，依托先进的语音技术，使用深度学习算法，将文本转换为自然流畅的语音。用户通过实时访问和调用API获取语音合成结果，将用户输入的文字合成为音频。通过音色选择、自定义音量、语速、音高等，可自定义音频格式，为企业和个人提供个性化的发音服务。

表 1-5 热词管理接口说明

接口类型	说明
热词管理	如果在您的业务领域有一些特有的词，默认识别效果差的时候可以考虑使用热词管理功能，将这些词添加到词表，改善识别效果。

1.2 调用说明

SIS服务提供了两种接口，包含REST（Representational State Transfer）API，支持您通过HTTPS请求调用，调用方法请参见**如何调用REST API**。也包含WebSocket接口，支持Websocket协议，调用方法请参见**如何调用WebSocket API**。

调用所需示例音频参见**示例音频**。

说明

调用接口的时候，无需开通服务，可直接调用。此时请按照实际需要选择计费方式，计费是按照调用接口的次数或者时长来计算费用，具体计费价格参见**语音交互价格计算器**。

支持两种计费方式：

- 按需计费，默认计费方式为“按需计费”。
- 折扣套餐包方式，是用户可以购买套餐包，扣费时调用次数会先在套餐包内进行抵扣，抵扣完后的剩余调用量默认转回按需计费方式。

1.3 终端节点

终端节点即调用API的请求地址，不同服务不同区域的终端节点不同。

目前语音交互服务一句话识别、录音文件识别支持以下地区和终端节点：

版权所有 © 华为云计算技术有限公司

说明

当前服务仅支持北京和上海区域，后续会陆续上线其他区域。

- 华北-北京四，推荐的区域，支持一句话识别、录音文件识别、实时语音识别、语音合成和热词等接口。
- 华东-上海一，推荐的区域，支持一句话识别、录音文件识别、实时语音识别、语音合成和热词等接口 。

表 1-6 一句话识别、录音文件识别

区域名称	区域	终端节点（Endpoint）	协议类型
华北-北京四	cn-north-4	sis-ext.cn-north-4.myhuaweicloud.com sis-ext.cn-north-4.myhuaweicloud.cn	https
华东-上海一	cn-east-3	sis-ext.cn-east-3.myhuaweicloud.com sis-ext.cn-east-3.myhuaweicloud.cn	https

表 1-7 实时语音识别

区域名称	区域	终端节点（Endpoint）	协议类型
华北-北京四	cn-north-4	sis-ext.cn-north-4.myhuaweicloud.com sis-ext.cn-north-4.myhuaweicloud.cn	Websocket
华东-上海一	cn-east-3	sis-ext.cn-east-3.myhuaweicloud.com sis-ext.cn-east-3.myhuaweicloud.cn	Websocket

语音合成支持的终端节点：

表 1-8 语音合成

区域名称	区域	终端节点（Endpoint）	协议类型
华北-北京四	cn-north-4	sis-ext.cn-north-4.myhuaweicloud.com sis-ext.cn-north-4.myhuaweicloud.cn	https

版权所有 © 华为云计算技术有限公司

区域名称	区域	终端节点（Endpoint）	协议类型
华东-上海一	cn-east-3	sis-ext.cn-east-3.myhuaweicloud.com sis-ext.cn-east-3.myhuaweicloud.cn	https

实时语音合成服务支持的终端节点：

表 1-9 实时语音合成

区域名称	区域	终端节点（Endpoint）	协议类型
华北-北京四	cn-north-4	sis-ext.cn-north-4.myhuaweicloud.com sis-ext.cn-north-4.myhuaweicloud.cn	Websocket
华东-上海一	cn-east-3	sis-ext.cn-east-3.myhuaweicloud.com sis-ext.cn-east-3.myhuaweicloud.cn	Websocket

热词（一句话识别、录音文件识别、实时语音识别中的热词）支持服务终端：

表 1-10 热词管理

区域名称	区域	终端节点（Endpoint）	协议类型
华北-北京四	cn-north-4	sis-ext.cn-north-4.myhuaweicloud.com sis-ext.cn-north-4.myhuaweicloud.cn	https
华东-上海一	cn-east-3	sis-ext.cn-east-3.myhuaweicloud.com sis-ext.cn-east-3.myhuaweicloud.cn	https

1.4 基本概念

- 账号

  用户注册华为云时的账号，账号对其所拥有的资源及云服务具有完全的访问权限，可以重置用户密码、分配用户权限等。由于账号是付费主体，为了确保账号安全，建议您不要直接使用账号进行日常管理工作，而是创建用户并使用他们进行日常管理工作。

版权所有 © 华为云计算技术有限公司

- 用户

  由账号在IAM中创建的用户，是云服务的使用人员，具有身份凭证（密码和访问密钥）。

  在**我的凭证**下，您可以查看账号ID和用户ID。通常在调用API的鉴权过程中，您需要用到账号、用户和密码等信息。

- 区域（Region）

  从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。

  详情请参见**区域和可用区**。

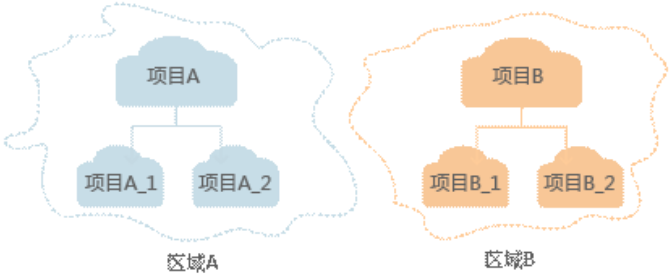
- 可用区（AZ，Availability Zone）

  一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

- 项目

  华为云的区域默认对应一个项目，这个项目由系统预置，用来隔离物理区域间的资源（计算资源、存储资源和网络资源），以默认项目为单位进行授权，用户可以访问您账号中该区域的所有资源。如果您希望进行更加精细的权限控制，可以在区域默认的项目中创建子项目，并在子项目中购买资源，然后以子项目为单位进行授权，使得用户仅能访问特定子项目中资源，使得资源的权限控制更加精确。

  图 1-1 项目隔离模型



版权所有 © 华为云计算技术有限公司

2 API 概览

SIS服务提供了两种接口，包含REST（Representational State Transfer）API，支持您通过HTTPS请求调用，请参见表2-1。也包含WebSocket接口，支持Websocket协议，请参见表2-2。

表 2-1 REST API 功能

接口	功能	API URI
一句话识别接口	一句话识别	POST /v1/{project_id}/asr/short-audio
录音文件识别接	录音文件识别接口	识别接口：POST /v1/{project_id}/asr/transcriber/jobs 状态查询：GET /v1/{project_id}/asr/transcriber/jobs/{job_id}
语音合成接口	语音合成接口	POST /v1/{project_id}/tts
热词管理	热词管理接口	创建热词表：POST /v1/{project_id}/asr/vocabularies 更新热词表：PUT /v1/{project_id}/asr/vocabularies/{vocabulary_id} 查询热词表信息：GET /v1/{project_id}/asr/vocabularies/{vocabulary_id} 删除热词表：DELETE /v1/{project_id}/asr/vocabularies/{vocabulary_id} 查询热词表列表：GET /v1/{project_id}/asr/vocabularies

版权所有 © 华为云计算技术有限公司

表 2-2 WebSocket API 功能

接口	功能	API URI
实时语音识别接口（请求）	开始识别	支持三种模式的请求消息： • **流式一句话** WSS /v1/{project_id}/rasr/short-stream • **实时语音识别连续模式** WSS /v1/{project_id}/rasr/continue-stream • **实时语音识别单句模式** WSS /v1/{project_id}/rasr/sentence-stream
	发送音频数据	
	结束识别	
实时语音合成（请求）	**Websocket握手请求**	WSS /v1/{project_id}/rtts

版权所有 © 华为云计算技术有限公司

3 如何调用 REST API

3.1 构造请求

本节介绍REST API请求的组成，并以调用一句话识别接口说明如何调用API。

您还可以通过这个视频教程了解如何构造请求调用API：**https://bbs.huaweicloud.com/videos/102987** 。

请求 URI

请求URI由如下部分组成。

{URI-scheme} :// {Endpoint} / {resource-path} ? {query-string}

表 3-1 请求 URI

参数	说明
URI-scheme	传输请求的协议，当前所有API均采用**HTTPS**协议。
Endpoint	承载REST服务端点的服务器域名或IP，不同服务在不同区域，Endpoint不同，可以从**终端节点**中获取。例如，一句话识别服务在“华北-北京四”区域的Endpoint为“sis-ext.cn-north-4.myhuaweicloud.com”。
resource-path	资源路径，即API访问路径。从具体API的URI模块获取，例如“一句话识别”API的resource-path为的“/v1/{project_id}/asr/short-audio”。其中“project_id”需要替换成用户的项目ID，可参考**获取项目ID**。
query-string	查询参数，可选，查询参数前面需要带一个“？”，形式为“参数名=参数取值”。例如“录音文件识别状态查询” 中“job_id=123”表示查询“job_id”为123的任务结果。

例如，在“华北-北京四””区域调用**一句话识别**API，则需要使用“华北-北京四”区域的Endpoint（sis-ext.cn-north-4.myhuaweicloud.com），拼接起来如下所示。

版权所有 © 华为云计算技术有限公司

```
https://sis-ext.cn-north-4.myhuaweicloud.com/v1/{project_id}/asr/short-audio
```

图 3-1 URI 示意图

说明

为查看方便，每个具体API的URI，只给出resource-path部分，并将请求方法写在一起。这是因为URI-scheme都是HTTPS，而Endpoint在同一个区域也相同，所以简洁起见将这两部分省略。

请求方法

HTTP请求方法（也称为操作或动词），它告诉服务你正在请求什么类型的操作。

- **GET**：请求服务器返回指定资源。
- **PUT**：请求服务器更新指定资源。
- **POST**：请求服务器新增资源或执行特殊操作。
- **DELETE**：请求服务器删除指定资源，如删除对象等。
- **HEAD**：请求服务器资源头部。
- **PATCH**：请求服务器更新资源的部分内容。当资源不存在的时候，PATCH可能会去创建一个新的资源。

在一句话识别的URI部分，您可以看到其请求方法为“POST”，则其请求为：

```
POST https://sis-ext.cn-north-4.myhuaweicloud.com/v1/{project_id}/asr/short-audio
```

请求消息头

附加请求头字段，如指定的URI和HTTP方法所要求的字段。例如定义消息体类型的请求头“Content-Type”，请求鉴权信息等。

如下公共消息头需要添加到请求中。

- **Content-Type**：消息体的类型（格式），必选，默认取值为“application/json”。
- **X-Auth-Token**：用户Token，可选，当使用Token方式认证时，必须填充该字段。用户Token请参考认证鉴权中的“Token认证”。

添加消息头后的请求如下所示：

```
POST https://sis-ext.cn-north-4.myhuaweicloud.com/v1/{project_id}/asr/short-audio
Content-Type: application/json
X-Auth-Token: MIINRwYJKoZIhvcNAQcCoIINOD...
```

请求消息体

请求消息体通常以结构化格式发出，与请求消息头中Content-type对应，传递除请求消息头之外的内容。若请求消息体中参数支持中文，则中文字符必须为UTF-8编码。

版权所有 © 华为云计算技术有限公司

每个接口的请求消息体内容不同，也并不是每个接口都需要有请求消息体（或者说消息体为空），GET、DELETE操作类型的接口就不需要消息体，消息体具体内容需要根据具体接口而定。

对于**一句话识别**接口，您可以从接口的请求部分看到所需的请求参数及参数说明。将消息体加入后的请求如下所示，其中“data”参数表示将音频转化为Base64编码字符串。

```
POST  https://sis-ext.cn-north-4.myhuaweicloud.com/v1/{project_id}/asr/short-audio
Content-Type: application/json
X-Auth-Token: MIINRwYJKoZIhvcNAQcCoIINOD...

{
    “data”: “encode audio by Base64”,
    “config”: {
        “audio_format”: “wav”,
        “property”: “chinese_8k_common”

    }
}
```

到这里为止这个请求需要的内容就具备齐全了，您可以使用**curl**、**Postman**或直接编写代码等方式发送请求调用API。对于**一句话识别**接口，您可以从响应消息部分看到返回参数及参数说明。

3.2 认证鉴权

调用接口有如下两种认证方式，您可以选择其中一种进行认证鉴权。

- Token认证：通过Token认证通用请求。
- AK/SK认证：通过AK（Access Key ID）/SK（Secret Access Key)加密调用请求。

Token 认证

说明

Token的有效期为24小时，需要使用同一个Token鉴权时，可以缓存起来，避免频繁调用。

Token在计算机系统中代表令牌（临时）的意思，拥有Token就代表拥有某种权限。Token认证就是在调用API的时候将Token加到请求消息头，从而通过身份认证，获得操作API的权限。

代码中加粗的斜体字段需要根据实际值填写***username***、***domainname***、***project name***，可登录控制台**“我的凭证”**页面获取。***password***为用户密码。

- 伪码

```
POST https://iam.cn-north-4.myhuaweicloud.com/v3/auth/tokens
Content-Type: application/json

{
    "auth": {
        "identity": {
```

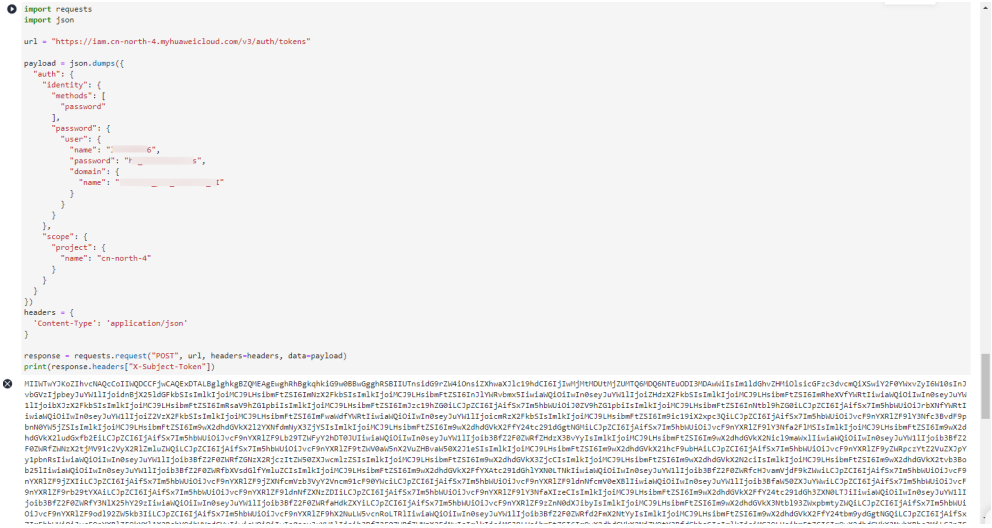

版权所有 © 华为云计算技术有限公司

```
            "methods": [
                "password"
            ],
            "password": {
                "user": {
                    "name": "username",
                    "password": "********",
                    "domain": {
                        "name": "domainname"
                    }
                }
            }
        },
        "scope": {
            "project": {
                "name": "cn-north-4"
            }
        }
    }
}
```

- Python

```
import requests
import json

url = "https://iam.cn-north-4.myhuaweicloud.com/v3/auth/tokens"
payload = json.dumps({
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "username",
          "password": "********",
          "domain": {
            "name": "domainname"
          }
        }
      }
    },
    "scope": {
      "project": {
        "name": "cn-north-4"
      }
    }
  }
})
headers = {
  'Content-Type': 'application/json'
}

response = requests.request("POST", url, headers=headers, data=payload)

print(response.headers["X-Subject-Token"])
```

图 3-2 获取 Token 示例



```
import requests
import json

url = "https://iam.cn-north-4.myhuaweicloud.com/v3/auth/tokens"

payload = json.dumps({
  "auth": {
    "identity": {
      "methods": [
        "password"
      ],
      "password": {
        "user": {
          "name": "[redacted]6",
          "password": "[redacted]s",
          "domain": {
            "name": "[redacted]"
          }
        }
      }
    },
    "scope": {
      "project": {
        "name": "cn-north-4"
      }
    }
  }
})
headers = {
  'Content-Type': 'application/json'
}

response = requests.request("POST", url, headers=headers, data=payload)
print(response.headers["X-Subject-Token"])
```

获取Token后，再调用其他接口时，您需要在请求消息头中添加“X-Auth-Token”的值。例如Token值为“ABCDEFJ....”，则调用接口时将“X-Auth-Token: ABCDEFJ....”加到请求消息头即可，如下所示。

```
Content-Type: application/json
X-Auth-Token: ABCDEFJ....
```

您还可以通过这个视频教程了解如何使用Token认证：**https://bbs.huaweicloud.com/videos/101333** 。

AK/SK 认证

说明

AK/SK签名认证方式仅支持消息体在12M以内，12M以上的请求请使用Token认证。

AK/SK认证就是使用AK/SK对请求进行签名，在请求时将签名信息添加到消息头，从而通过身份认证。

- AK(Access Key ID)：访问密钥ID。与私有访问密钥关联的唯一标识符；访问密钥ID和私有访问密钥一起使用，对请求进行加密签名。
- SK(Secret Access Key)：与访问密钥ID结合使用的密钥，对请求进行加密签名，可标识发送方，并防止请求被修改。

使用AK/SK认证时，您可以基于签名算法使用AK/SK对请求进行签名，也可以使用专门的签名SDK对请求进行签名。详细的签名方法和SDK使用方法请参见**API签名指南**。

须知

签名SDK只提供签名功能，与服务提供的SDK不同，使用时请注意。

AK/SK获取方式请参考**获取AK/SK**。

版权所有 © 华为云计算技术有限公司

3.3 返回结果

状态码

请求发送以后，您会收到响应，包含状态码、响应消息头和消息体。

状态码是一组从1xx到5xx的数字代码，状态码表示了请求响应的状态，完整的状态码列表请参见**状态码**。

对于SIS服务接口，如果调用后返回状态码为“200”，则表示请求成功。

响应消息头

对应请求消息头，响应同样也有消息头，如“Content-type”。SIS服务响应消息头无特殊用途，可用于定位问题使用。

响应消息体

响应消息体通常以结构化格式返回，与响应消息头中Content-type对应，传递除响应消息头之外的内容。

对于**一句话识别**接口，返回如下消息体。格式请具体参考**一句话识别**响应消息部分。

```
{
  "trace_id": "567e8537-a89c-13c3-a882-826321939651",
  "result":{
    "text": "欢迎使用语音云服务。",
    "score": 0.9
  }
}
```

当接口调用出错时，会返回错误码及错误信息说明，错误响应的Body体格式如下所示。

```
{
    "error_msg": "***",
    "error_code": "SIS.0001"
}
```

其中，error_code表示错误码，error_msg表示错误描述信息。

版权所有 © 华为云计算技术有限公司

4 如何调用 WebSocket API

前提条件

在调用实时语音识别的Websocket接口之前，您需要完成Token认证，详细操作指导请参见认证鉴权。

接口使用介绍

实时语音识别接口使用Websocket协议承载，客户端与服务端交流流程如图4-1所示。

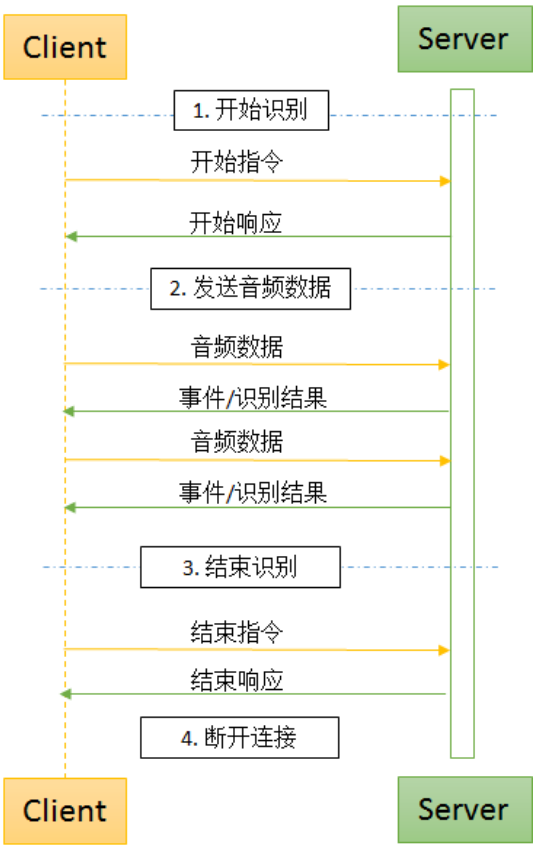
分为三个主要步骤：

1. WebSocket握手。
2. 基于WebSocket协议进行实时语音识别。
3. 关闭WebSocket连接。

其中，基于WebSocket协议进行实时语音识别时，需要客户端首先发送转写开始消息，然后持续发送语音数据至服务端，最后发送结束消息，在此期间客户端会持续收到服务端发送的转写结果或事件，客户端根据所收到的响应消息做对应处理。实时语音识别接口的具体细节请参见接口说明章节。

版权所有 © 华为云计算技术有限公司

图 4-1 客户端与服务端交流流程



版权所有 © 华为云计算技术有限公司

5 实时语音识别接口

5.1 接口说明

实时语音识别接口基于Websocket协议实现。分别提供了“流式一句话”、“实时语音识别连续模式”、“实时语音识别单句模式”三种模式。

三种模式的握手请求wss-URI不同，基于Websocket承载的实时语音识别请求和响应的消息格式相同。

开发者可以使用java、python、c++等开发语言支持Websocket的对应软件包或库接口，与实时语音识别引擎握手连接，并发送语音数据和接收转写结果，最后关闭Websocket连接。

- 获取Websocket握手请求wss-URI请参见**Websocket握手请求**。
- 获取实时语音识别请求消息格式信息请参见**实时语音识别请求**。
- 获取实时语音识别响应消息格式请参见**实时语音识别响应**。

客户端通过Websocket协议访问实时流转写接口时，连接时长不能超过5小时。超过5小时的Websocket连接，服务端会自动断链。

5.2 Websocket 握手请求

5.2.1 流式一句话

功能介绍

流式一句话模式的语音长度限制为一分钟，适合于对话聊天等识别场景。

该接口支持用户将一整段语音分段，以流式输入，最后得到识别结果。实时语音识别引擎在获得分段的输入语音的同时，就可以同步地对这段数据进行特征提取和解码工作，而不用等到所有数据都获得后再开始工作。因此这样就可以在最后一段语音结束后，仅延迟很短的时间（也即等待处理最后一段语音数据以及获取最终结果的时间）即可返回最终识别结果。这种流式输入方式能缩短整体上获得最终结果的时间，极大地提升用户体验。

版权所有 © 华为云计算技术有限公司

wss-URI

- wss-URI格式

  wss /v1/{project_id}/rasr/short-stream

- 参数说明

表 5-1 参数说明

参数名	是否必选	说明
project_id	是	项目编号。获取方法，请参见**获取项目ID**。

表 5-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“**企业项目管理**”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 • 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 • 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 • 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

- 请求示例（伪码）

```
wss://{endpoint}/v1/{project_id}/rasr/short-stream

Request Header:
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```


版权所有 © 华为云计算技术有限公司

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- Python3语言请求代码示例

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装websocket-client, 执行pip install websocket-client
import websocket
import threading
import time
import json

def rasr_demo():
    url = 'wss://{{endpoint}}/v1/{{project_id}}/rasr/short-stream'  # endpoint和project_id需替换
    audio_path = '音频路径'
    token = '用户对应region的token'
    header = {
        'X-Auth-Token': token
    }
    with open(audio_path, 'rb') as f:
        data = f.read()
    body = {
        'command': 'START',
        'config': {
            'audio_format': 'pcm8k16bit',
            'property': 'chinese_8k_general'
        }
    }
    def _on_message(ws, message):
        print(message)
    def _on_error(ws, error):
        print(error)
    ws = websocket.WebSocketApp(url, header, on_message=_on_message, on_error=_on_error)
    _thread = threading.Thread(target=ws.run_forever, args=(None, None, 30, 20))
    _thread.start()
    time.sleep(1)
    ws.send(json.dumps(body), opcode=websocket.ABNF.OPCODE_TEXT)
    now_index = 0
    byte_len = 4000
    while now_index < len(data):
        next_index = now_index + byte_len
        if next_index > len(data):
            next_index = len(data)
        send_array = data[now_index: next_index]
        ws.send(send_array, opcode=websocket.ABNF.OPCODE_BINARY)
        now_index += byte_len
        time.sleep(0.05)
    ws.send("{\"command\": \"END\", \"cancel\": \"false\"}", opcode=websocket.ABNF.OPCODE_TEXT)
    time.sleep(10)
    ws.close()
if __name__ == '__main__':
    rasr_demo()
```

- Java语言请求代码示例

```
import okhttp3.OkHttpClient;
import okhttp3.Request;
import okhttp3.Response;
import okhttp3.WebSocket;
import okhttp3.WebSocketListener;
import okio.ByteString;
import java.net.URL;
/**
 * 此demo仅供测试使用，强烈建议使用SDK
 * 使用前需已配置okhttp、okio jar包。jar包可通过下载SDK获取。
 */
public class RasrDemo {
  public void rasrDemo() {
```


版权所有 © 华为云计算技术有限公司

```
    try {
      // endpoint和projectId需要替换成实际信息。
      String url = "wss://{{endpoint}}/v1/{{project_id}}/rasr/short-stream";
      String token = "对应region的token";
      byte[] data = null;  // 存放将要发送音频的byte数组
      OkHttpClient okHttpClient = new OkHttpClient();
      Request request = new Request.Builder().url(url).header("X-Auth-Token", token).build();
      WebSocket webSocket = okHttpClient.newWebSocket(request, new MyListener());
      webSocket.send("{\"command\": \"START\", \"config\": {\"audio_format\": \"pcm8k16bit\",
\"property\": \"chinese_8k_general\"}}");
      webSocket.send(ByteString.of(data));
      webSocket.send("{  \"command\": \"END\",  \"cancel\": false}");
      Thread.sleep(10000);
      webSocket.close(1000, null);

    } catch (Exception e) {
      e.printStackTrace();
    }


  }

  class MyListener extends WebSocketListener {
    @Override

    public void onOpen(WebSocket webSocket, Response response) {
      System.out.println("conneected");
    }
    @Override
    public void onClosed(WebSocket webSocket, int code, String reason) {
      System.out.println("closed");
    }
    @Override
    public void onFailure(WebSocket webSocket, Throwable t, Response response) {
      t.printStackTrace();
    }
    @Override
    public void onMessage(WebSocket webSocket, String text) {
      System.out.println(text);
    }
  }
  public static void main(String[] args) {
    RasrDemo rasrDemo = new RasrDemo();
    rasrDemo.rasrDemo();
  }
}
```

- JavaScript（nodejs推荐使用v18.20.2 (LTS)版本）

```
// 导入 Node.js 的 ws 库
const WebSocket = require('ws');
function shortStreamDemo(endpoint,audioPath, projectID, token) {
    const url = `wss://${endpoint}/v1/${projectID}/rasr/short-stream`;  // 替换 endpoint 和 projectID
    // 读取音频文件内容
    const fs = require('fs');
    let data = fs.readFileSync(audioPath);
    // HTTP Headers中携带Token
    const headers = {
        'X-Auth-Token': token,
        // 企业id 可选加
        // 'Enterprise-Project-Id': 企业id
    };
    // 创建WebSocket实例
    const ws = new WebSocket(url, {
        headers // 添加自定义HTTP头
    });
    ws.on('open', async () => {
        const body = {
            command: 'START',
            config: {
                audio_format: 'pcm16k16bit',
```


版权所有 © 华为云计算技术有限公司

```
            property: 'chinese_16k_general'
        }
    };
    ws.send(JSON.stringify(body));
    let nowIndex = 0;
    const byteLen = 3200; // 禁止空值输入 建议范围 2000-10000
    while (nowIndex < data.length) {
        const nextIndex = nowIndex + byteLen;
        const sendArray = data.slice(nowIndex, nextIndex > data.length ? data.length : nextIndex);
        ws.send(sendArray, { binary: true });
        nowIndex += byteLen;
        await new Promise(resolve => setTimeout(resolve, 100)); // 模拟延时(单位ms)
    }
    const endCommand = JSON.stringify({ command: 'END', cancel: 'false' });
    ws.send(endCommand);
  });
  ws.on('message', (data) => {
    if (data instanceof Buffer) {
        // 将Buffer转换为UTF-8编码的字符串
        const messageString = data.toString('utf8');
        console.log('Received (converted from Buffer):', messageString);
        const type = JSON.parse(messageString).resp_type;
        if (type ==='END'|| type ==='ERROR') {
            ws.close();
        }
    }
  });
  ws.on('error', (error) => {
    console.error('WebSocket Error:', error);
  });
};
shortStreamDemo(endpoint,audioPath, projectID, token);
```

5.2.2 实时语音识别连续模式

功能介绍

连续识别模式的语音总长度限制为五小时，适合于会议、演讲和直播等场景。

连续识别模式在流式识别的基础上，结合了语音的端点检测功能。语音数据也是分段输入，但是连续识别模式将会在处理数据之前进行端点检测，如果是语音才会进行实际的解码工作，如果检测到静音，将直接丢弃。如果检测到一段语音的结束点，就会直接将当前这一段的识别结果返回，然后继续检测后面的语音数据。因此在连续识别模式中，可能多次返回识别结果。如果送入的一段语音较长，甚至有可能在一次返回中包括了多段的识别结果。

由于引入了静音检测，连续识别模式通常会比流式识别能具有更高的效率，因为对于静音段将不会进行特征提取和解码操作，因而能更有效地利用CPU。而流式识别通常和客户端的端点检测功能相结合，只将检测到的有效语音段上传到服务器进行识别。

wss-URI

- wss-URI格式

  wss /v1/{project_id}/rasr/continue-stream

- 参数说明

版权所有 © 华为云计算技术有限公司

表 5-3 参数说明

参数名	是否必选	说明
project_id	是	项目编号。获取方法，请参见**获取项目ID**。

表 5-4 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“**企业项目管理**”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

- 请求示例（伪码）

```
wss://{endpoint}/v1/{project_id}/rasr/continue-stream

Request Header:
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- Python3语言请求代码示例

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装websocket-client, 执行pip install websocket-client
```


版权所有 © 华为云计算技术有限公司

```
import websocket
import threading
import time
import json

def rasr_demo():
    url = 'wss://{{endpoint}}/v1/{{project_id}}/rasr/continue-stream'  # endpoint和project_id需替换
    audio_path = '音频路径'
    token = '用户对应region的token'
    header = {
        'X-Auth-Token': token
    }
    with open(audio_path, 'rb') as f:
        data = f.read()
    body = {
        'command': 'START',
        'config': {
            'audio_format': 'pcm8k16bit',
            'property': 'chinese_8k_general'
        }
    }
    def _on_message(ws, message):
        print(message)
    def _on_error(ws, error):
        print(error)
    ws = websocket.WebSocketApp(url, header, on_message=_on_message, on_error=_on_error)
    _thread = threading.Thread(target=ws.run_forever, args=(None, None, 30, 20))
    _thread.start()
    time.sleep(1)
    ws.send(json.dumps(body), opcode=websocket.ABNF.OPCODE_TEXT)
    now_index = 0
    byte_len = 4000
    while now_index < len(data):
        next_index = now_index + byte_len
        if next_index > len(data):
            next_index = len(data)
        send_array = data[now_index: next_index]
        ws.send(send_array, opcode=websocket.ABNF.OPCODE_BINARY)
        now_index += byte_len
        time.sleep(0.05)
    ws.send("{\"command\": \"END\", \"cancel\": \"false\"}", opcode=websocket.ABNF.OPCODE_TEXT)
    time.sleep(10)
    ws.close()
if __name__ == '__main__':
    rasr_demo()
```

- Java语言请求代码示例

```
import okhttp3.OkHttpClient;
import okhttp3.Request;
import okhttp3.Response;
import okhttp3.WebSocket;
import okhttp3.WebSocketListener;
import okio.ByteString;
import java.net.URL;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 * 使用前需已配置okhttp、okio jar包。jar包可通过下载SDK获取。
 */
public class RasrDemo {
  public void rasrDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      String url = "wss://{{endpoint}}/v1/{{project_id}}/rasr/continue-stream";
      String token = "对应region的token";
      byte[] data = null;  // 存放将要发送音频的byte数组
      OkHttpClient okHttpClient = new OkHttpClient();
      Request request = new Request.Builder().url(url).header("X-Auth-Token", token).build();
      WebSocket webSocket = okHttpClient.newWebSocket(request, new MyListener());
      webSocket.send("{\"command\": \"START\", \"config\": {\"audio_format\": \"pcm8k16bit\",
```


版权所有 © 华为云计算技术有限公司

```
\"property\": \"chinese_8k_general\"}});
      webSocket.send(ByteString.of(data));
      webSocket.send("{  \"command\": \"END\",  \"cancel\": false}");
      Thread.sleep(10000);
      webSocket.close(1000, null);
    } catch (Exception e) {
      e.printStackTrace();
    }
  }
  class MyListener extends WebSocketListener {
    @Override
    public void onOpen(WebSocket webSocket, Response response) {
      System.out.println("conneected");
    }
    @Override
    public void onClosed(WebSocket webSocket, int code, String reason) {
      System.out.println("closed");
    }
    @Override
    public void onFailure(WebSocket webSocket, Throwable t, Response response) {
      t.printStackTrace();
    }
    @Override
    public void onMessage(WebSocket webSocket, String text) {
      System.out.println(text);
    }

  }
  public static void main(String[] args) {
    RasrDemo rasrDemo = new RasrDemo();
    rasrDemo.rasrDemo();
  }
}
```

- JavaScript（nodejs推荐使用v18.20.2 (LTS)版本）

```
// 导入 Node.js 的 ws 库
const WebSocket = require('ws');
function continueStremDemo(endpoint,audioPath, projectID, token) {
    const url = `wss://${endpoint}/v1/${projectID}/rasr/continue-stream`;  // 替换 endpoint 和 projectID
    // 读取音频文件内容
    const fs = require('fs');
    let data = fs.readFileSync(audioPath);
    // HTTP Headers中携带Token
    const headers = {
        'X-Auth-Token': token,
        // 企业id 可选加
        // 'Enterprise-Project-Id': 企业id
    };
    // 创建WebSocket实例
    const ws = new WebSocket(url, {
        headers // 添加自定义HTTP头
    });
    ws.on('open', async () => {
        const body = {
            command: 'START',
            config: {
                audio_format: 'pcm16k16bit',
                property: 'chinese_16k_general'
            }
        };
        ws.send(JSON.stringify(body));
        let nowIndex = 0;
        const byteLen = 3200; // 禁止空值输入 建议范围 2000-10000
        while (nowIndex < data.length) {
            const nextIndex = nowIndex + byteLen;
            const sendArray = data.slice(nowIndex, nextIndex > data.length ? data.length : nextIndex);
            ws.send(sendArray, { binary: true });
            nowIndex += byteLen;
            await new Promise(resolve => setTimeout(resolve, 100)); // 模拟延时(单位ms)
        }
```


版权所有 © 华为云计算技术有限公司

```
        const endCommand = JSON.stringify({ command: 'END', cancel: 'false' });
        ws.send(endCommand);
    });
    ws.on('message', (data) => {
        if (data instanceof Buffer) {
            // 将Buffer转换为UTF-8编码的字符串
            const messageString = data.toString('utf8');
            console.log('Received (converted from Buffer):', messageString);
            const type = JSON.parse(messageString).resp_type;
            if (type ==='END' || type ==='ERROR') {
                ws.close();
            }
        }
    });
    ws.on('error', (error) => {
        console.error('WebSocket Error:', error);
    });
};
continueStremDemo(endpoint,audioPath, projectID, token);
```

5.2.3 实时语音识别单句模式

功能介绍

单句模式自动检测一句话的结束，因此适合于需要与您的系统进行交互的场景，例如外呼、控制口令等场景。

实时语音识别引擎的单句识别模式，和连续识别模式类似，也会进行语音的端点检测，如果检测到静音，将直接丢弃，检测到语音才会馈入核心进行实际的解码工作，如果检测到一段语音的结束点，就会将当前这一段的识别结果返回。和连续识别不同的是，在单句模式下，返回第一段的识别结果后，将不再继续识别后续的音频。这主要是用于和用户进行语音交互的场景下，当用户说完一句话后，往往会等待后续的交互操作，例如聆听根据识别结果播报的相关内容，因而没有必要继续识别后续的音频。

wss-URI

- wss-URI格式：

  wss /v1/{project_id}/rasr/sentence-stream

- 参数说明

表 5-5 参数说明

参数名	是否必选	说明
project_id	是	项目编号。获取方法，请参见**获取项目ID**。

表 5-6 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“企业项目管理”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

- 请求示例（伪码）

```
wss://{endpoint}/v1/{project_id}/rasr/sentence-stream

Request Header:
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见终端节点。

- Python3语言请求代码示例

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装websocket-client, 执行pip install websocket-client
import websocket
import threading
import time
import json

def rasr_demo():
    url = 'wss://{{endpoint}}/v1/{{project_id}}/rasr/sentence-stream'  # endpoint和project_id需替换
    audio_path = '音频路径'
    token = '用户对应region的token'
    header = {
        'X-Auth-Token': token
    }
    with open(audio_path, 'rb') as f:
        data = f.read()
    body = {
        'command': 'START',
        'config': {
            'audio_format': 'pcm8k16bit',
```


版权所有 © 华为云计算技术有限公司

```
            'property': 'chinese_8k_general'
        }
    }
    def _on_message(ws, message):
        print(message)
    def _on_error(ws, error):
        print(error)
    ws = websocket.WebSocketApp(url, header, on_message=_on_message, on_error=_on_error)
    _thread = threading.Thread(target=ws.run_forever, args=(None, None, 30, 20))
    _thread.start()
    time.sleep(1)
    ws.send(json.dumps(body), opcode=websocket.ABNF.OPCODE_TEXT)
    now_index = 0
    byte_len = 4000
    while now_index < len(data):
        next_index = now_index + byte_len
        if next_index > len(data):
            next_index = len(data)
        send_array = data[now_index: next_index]
        ws.send(send_array, opcode=websocket.ABNF.OPCODE_BINARY)
        now_index += byte_len
        time.sleep(0.05)
    ws.send("{\"command\": \"END\", \"cancel\": \"false\"}", opcode=websocket.ABNF.OPCODE_TEXT)
    time.sleep(10)
    ws.close()
if __name__ == '__main__':
    rasr_demo()
```

- Java语言请求代码示例

```
import okhttp3.OkHttpClient;
import okhttp3.Request;
import okhttp3.Response;
import okhttp3.WebSocket;
import okhttp3.WebSocketListener;
import okio.ByteString;
import java.net.URL;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 * 使用前需已配置okhttp、okio jar包。jar包可通过下载SDK获取。
 */
public class RasrDemo {
  public void rasrDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      String url = "wss://{{endpoint}}/v1/{{project_id}}/rasr/sentence-stream";
      String token = "对应region的token";
      byte[] data = null;  // 存放将要发送音频的byte数组
      OkHttpClient okHttpClient = new OkHttpClient();
      Request request = new Request.Builder().url(url).header("X-Auth-Token", token).build();
      WebSocket webSocket = okHttpClient.newWebSocket(request, new MyListener());
      webSocket.send("{\"command\": \"START\", \"config\": {\"audio_format\": \"pcm8k16bit\",
\"property\": \"chinese_8k_general\"}}");
      webSocket.send(ByteString.of(data));
      webSocket.send("{  \"command\": \"END\",  \"cancel\": false}");
      Thread.sleep(10000);
      webSocket.close(1000, null);
    } catch (Exception e) {
      e.printStackTrace();
    }
  }

  class MyListener extends WebSocketListener {
    @Override
    public void onOpen(WebSocket webSocket, Response response) {
      System.out.println("conneected");
    }
    @Override
    public void onClosed(WebSocket webSocket, int code, String reason) {
      System.out.println("closed");
```


版权所有 © 华为云计算技术有限公司

```
    }
    @Override
    public void onFailure(WebSocket webSocket, Throwable t, Response response) {
      t.printStackTrace();
    }
    @Override
    public void onMessage(WebSocket webSocket, String text) {
      System.out.println(text);
    }
  }
  public static void main(String[] args) {
    RasrDemo rasrDemo = new RasrDemo();
    rasrDemo.rasrDemo();
  }
}
```

- JavaScript（nodejs推荐使用v18.20.2 (LTS)版本）

```
// 导入 Node.js 的 ws 库
const WebSocket = require('ws');
function sentenceStreamDemo(endpoint,audioPath, projectID, token) {
    const url = `wss://${endpoint}/v1/${projectID}/rasr/sentence-stream`;  // 替换 endpoint 和 projectID
    // 读取音频文件内容
    const fs = require('fs');
    let data = fs.readFileSync(audioPath);
    // HTTP Headers中携带Token
    const headers = {
        'X-Auth-Token': token,
        //  企业id 可选加
        // 'Enterprise-Project-Id': 企业id
    };
    // 创建WebSocket实例
    const ws = new WebSocket(url, {
        headers // 添加自定义HTTP头
    });
    ws.on('open', async () => {
        const body = {
            command: 'START',
            config: {
                audio_format: 'pcm16k16bit',
                property: 'chinese_16k_general'
            }
        };
        ws.send(JSON.stringify(body));
        let nowIndex = 0;
        const byteLen = 3200; // 禁止空值输入 建议范围 2000-10000
        while (nowIndex < data.length) {
            const nextIndex = nowIndex + byteLen;
            const sendArray = data.slice(nowIndex, nextIndex > data.length ? data.length : nextIndex);
            ws.send(sendArray, { binary: true });
            nowIndex += byteLen;
            await new Promise(resolve => setTimeout(resolve, 100)); // 模拟延时(单位ms)
        }
        const endCommand = JSON.stringify({ command: 'END', cancel: 'false' });
        ws.send(endCommand);
    });
    ws.on('message', (data) => {
        if (data instanceof Buffer) {
            // 将Buffer转换为UTF-8编码的字符串
            const messageString = data.toString('utf8');
            console.log('Received (converted from Buffer):', messageString);
            const type = JSON.parse(messageString).resp_type;
            if (type ==='END'|| type ==='ERROR') {
                ws.close();
            }
        }
    });
    ws.on('error', (error) => {
        console.error('WebSocket Error:', error);
    });
```


版权所有 © 华为云计算技术有限公司

```
};
sentenceStreamDemo(endpoint,audioPath, projectID, token);
```

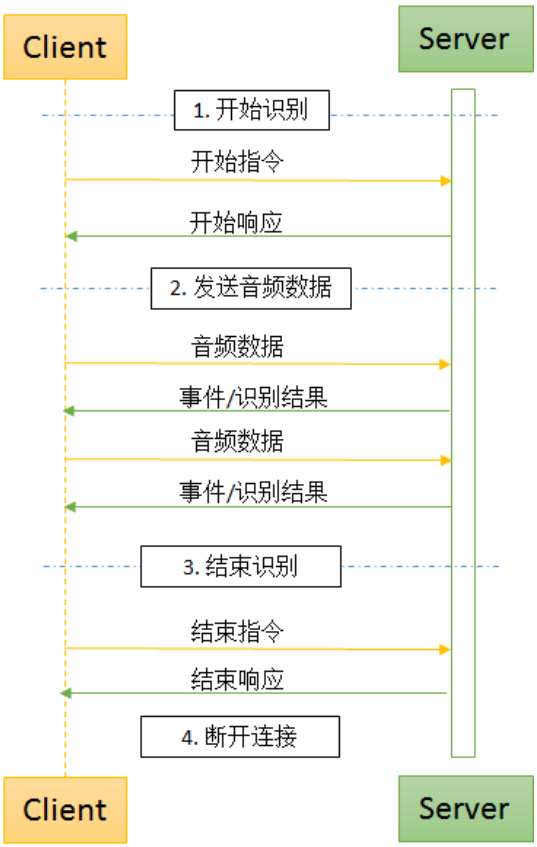
5.3 实时语音识别请求

5.3.1 实时语音识别工作流程

实时语音识别分为开始识别、发送音频数据、结束识别，断开连接四个阶段。

- 开始阶段需要发送开始指令，包含采样率，音频格式，是否返回中间结果等配置信息。服务端会返回一个开始响应。
- 发送音频阶段客户端会分片发送音频数据，服务会返回识别结果或者其他事件，如音频超时，静音部分过长等。
- 音频发送结束后，客户端会发送结束请求，服务端会返回end响应。
- 实时语音识别必须客户端要主动断开连接。当服务端超过20s没有收到客户的任何数据时，会返回error事件，并主动断开。

图 5-1 工作流程



版权所有 © 华为云计算技术有限公司

5.3.2 开始识别

功能介绍

当wss握手请求收到成功响应后，客户端到服务端的通信协议会升级为Websocket协议。通过Websocket协议，客户端发送开始识别请求，用于配置实时语音识别的配置信息。

请求消息

表 5-7 参数说明

参数名	是否必选	参数类型	说明
command	是	String	表示客户端发送开始识别请求，参数值需设置为START。
config	是	Object	配置信息。结构信息请参见**表 config 数据结构**。

表 5-8 config 数据结构

参数	是否必选	参数类型	说明
audio_format	是	String	支持语音的格式，请参见**表 audio_format取值范围**。
property	是	String	所使用的模型特征串。通常是“语种_采样率_领域”的形式，例如chinese_8k_common。请参见**表 property取值范围**。
add_punc	否	String	表示是否在识别结果中添加标点，取值为yes 、 no，默认no。
digit_norm	否	String	表示是否将语音中的数字识别为阿拉伯数字，取值为yes 、 no，默认为yes。
vad_head	否	Integer	一段音频的开头，当它的静音持续时间大于等于此值时，在实时语音识别单句模式下将返回“长时间静音”事件（EXCEEDED_SILENCE）并结束识别；在连续模式下将会断句并继续下一句的识别。在流式一句话模式下此参数不生效。 如果设置为0，等价于设置为60000。 取值范围：[0, 60000]的整数，单位为ms，默认为10000ms，即10s。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	说明
vad_tail	否	Integer	音频的结尾中的静音时间，正常情况下不应设成很小的值。 如果检测语音结尾的静音时长大于等于此值时，在实时语音识别单句模式下将返回VOICE_END（识别结果非空）或EXCEEDED_SILENCE（识别结果为空）事件并结束识别；在连续模式下将会断句并继续下一句的识别。在流式一句话模式下此参数不生效。 取值范围：[0, 3000]的整数，单位为ms，默认为500ms。 注意：vad_tail若设置过小（<200ms），会导致断句过于频繁，影响识别结果。
max_seconds	否	Integer	一句话的最大时长。如果检测到语音持续时间大于等于此值时，在实时语音识别单句模式下将返回VOICE_END（识别结果非空）或EXCEEDED_SILENCE（识别结果为空）事件并结束识别，在连续模式下将会断句并继续下一句的识别。在流式一句话模式下此参数不生效。 取值范围：[1, 60]的整数，单位为s，默认为30s。
interim_results	否	String	是否输出中间结果，可以为yes或no。默认为no，表示不输出中间结果。
vocabulary_id	否	String	热词表id，不使用热词则不填写。 创建热词表信息请参考**创建热词表**。
need_word_info	否	String	表示是否在识别结果中输出分词结果信息，取值为“yes”和“no”，默认为“no”。对于中文，是按照字级别进行分词；对于英文是按照子词（BPE）级别进行分词。

表 5-9 property 取值范围

property取值	说明
chinese_8k_general	支持采样率为8k的中文普通话语音识别，采用新一代端到端识别算法，识别准确率更高。 区域支持cn-east-3和cn-north-4（强烈推荐使用）。

版权所有 © 华为云计算技术有限公司

property取值	说明
chinese_16k_general	支持采样率为16k的中文普通话语音识别，采用新一代端到端识别算法，识别准确率更高。 区域支持cn-east-3和cn-north-4（强烈推荐使用）。
english_16k_general	支持采样率为16k的英语语音识别，采用新一代端到端识别算法，识别准确率更高。区域仅支持cn-north-4，不支持数字归一化（digit_norm参数）。
arabic_16k_general	支持采样率为16k的阿拉伯语语音识别，采用新一代端到端识别算法，兼容标准阿拉伯语、埃及方言、沙特方言、阿联酋方言。区域仅支持cn-north-4，不支持标点预测（add_punc参数）、数字归一化（digit_norm参数）和热词（vocabulary_id参数）功能。
sichuan_16k_common	支持采样率为16k的中文普通话与四川话方言语音识别。区域仅支持cn-north-4。
cantonese_16k_common	支持采样率为16k的粤语方言语音识别。区域仅支持cn-north-4。
shanghai_16k_common	支持采样率为16k的上海话方言语音识别。区域仅支持cn-north-4。
chinese_16k_court	支持采样率为16k的庭审会议语音识别。区域仅支持cn-north-4。
chinese_16k_it	支持采样率为16k的IT会议语音识别。区域仅支持cn-north-4。
chinese_8k_common	支持采样率为8k的中文普通话语音识别，此为旧版模型，建议使用chinese_8k_general。
chinese_16k_common	支持采样率为16k的中文普通话语音识别，此为旧版模型，建议使用chinese_16k_general。

表 5-10 audio_format 取值范围

audio_format取值	说明
pcm16k16bit	16k16bit单通道录音数据。
pcm8k16bit	8k16bit单通道录音数据。
ulaw16k8bit	16k8bit ulaw单通道录音数据。
ulaw8k8bit	8k8bit ulaw单通道录音数据。
alaw16k8bit	16k8bit alaw单通道录音数据。
alaw8k8bit	8k8bit alaw单通道录音数据。

版权所有 © 华为云计算技术有限公司

说明

目前仅支持裸音频格式，仅支持pcm编码的wav格式，不支其他wav头或者arm格式的编码。

示例

```
{
  "command": "START",
  "config":
  {
    "audio_format": "ulaw8k8bit",
    "property": "chinese_8k_general",
    "add_punc": "yes",
    "vad_tail": 400,
    "interim_results": "yes",
    "need_word_info": "yes"
  }
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.3.3 发送音频数据

在收到“开始识别”的响应之后，可以开始发送音频数据。为节省流量，音频以二进制数据帧形式（binary message）的方式发送。

音频数据将分片发送，也即在获得一定量音频数据的同时就可以发送一个binary message，每个分片建议在50ms~1000ms之间，建议在需要实时反馈的情况下100ms，不需要实时反馈的情况下500ms。实时语音识别代码示例请参考**SDK文档**。

当前SIS服务对于8k音频的分片大小限制为[160, 32768]字节, 16k音频的分片大小限制为[320, 65536]字节, 分片大小超出上限或低于下限会报错。

5.3.4 结束识别

功能介绍

对于识别中的对话，需要在Websocket上发送“结束识别”的请求来取消或结束识别。 "结束识别"请求使用文本类型的数据帧(text message)发送，命令和参数以json字符串的形式提供。

请求消息

表 5-11 参数说明

参数名	是否必选	参数类型	说明
command	是	String	表示客户端结束识别请求，参数值设置为END。

版权所有 © 华为云计算技术有限公司

参数名	是否必选	参数类型	说明
cancel	否	Boolen	是否取消返回识别结果。 ● true：表示取消识别，也即丢弃识别中和未识别的语音数据并结束，不返回剩余的识别结果。 ● false：表示继续处理识别中和未识别的语音数据直到处理完所有之前发送的数据。默认是false。

示例

```
{
  "command": "END",
  "cancel": false
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.4 实时语音识别响应

5.4.1 开始识别请求响应

由于WebSocket是全双工的，因此响应就是从服务器端发送给客户端的消息，但也并不是所有的请求信息都有一条对应的响应。服务器端收到“开始识别”请求时，会给出如下响应消息，以json字符串形式放置在text message中。

响应消息

表 5-12 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为START，表示开始识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。

示例

```
{
   "resp_type": "START",
```


版权所有 © 华为云计算技术有限公司

```
    "trace_id": "567e8537-a89c-13c3-a882-826321939651"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.4.2 事件响应

服务器端检测到某些事件时，会给出如下响应消息，以json字符串形式放置在text message中。

响应消息

表 5-13 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为EVENT，表示开始识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
event	String	具体的事件，详细参数可参见**event取值范围及其说明**。
timestamp	Integer	保留字段。将来会用于此事件发生的具体时间，以会话开始作为0点，单位为ms。

event 取值范围及其说明

表 5-14 event 取值范围

事件	说明
VOICE_START	检测到句子开始。
VOICE_END	检测到句子结束。
EXCEEDED_SILENCE	静音超长，也即没有检测到声音。

- **在流式一句话模式下：**
  - 不会返回VOICE_START、VOICE_END、EXCEEDED_SILCENCE事件。
- **在实时语音识别单句模式下：**

版权所有 © 华为云计算技术有限公司

  - 返回VOICE_START事件，表示检测到语音，此时IVR可以做打断。
  - 返回VOICE_END事件后，表示一句话结束，后续的音频将被忽略，不会再进行识别。
  - 只会返回最多一组VOICE_START和VOICE_END事件。
  - 如果返回EXCEEDED_SILENCE事件，表示超过vad_head没有检测到声音，通常表示用户一直没有说话。此时后续的音频将被忽略，不会再进行识别。
- **在实时语音识别连续模式下：**
  - 不会返回VOICE_START、VOICE_END、EXCEED_SLIENCE事件。

示例

```
{
  "resp_type": "EVENT",
  "trace_id": "567e8537-a89c-13c3-a882-826321939651",
  "event": "VOICE_END",
  "timestamp": 1500
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.4.3 识别结果响应

服务端在收到客户端发送的连续音频数据后， 当服务端识别出结果后会实时向客户端按句推送识别结果响应消息， 以json字符串形式放置在text message中。

响应消息

表 5-15 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为RESULT，表示识别结果响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
segments	Array of objects	多句结果。 请参考**表 segment 数据结构**。

版权所有 © 华为云计算技术有限公司

表 5-16 segment 数据结构

参数名	参数类型	说明
start_time	Integer	相对时间戳，表示一句的起始，单位为ms。
end_time	Integer	相对时间戳，表示一句的结束，单位为ms。
is_final	Boolen	输出是否为最终结果。 true表示是最终结果， false表示为中间临时结果。
result	Object	调用成功表示识别结果，调用失败时无此字段。 请参考**表 result数据结构**。

表 5-17 result 数据结构

参数名	参数类型	说明
text	String	识别结果。
score	Float	识别结果的置信度，取值范围：0~1。此值会在实时识别中不断刷新，直到返回最终结果。 **说明** 临时结果的置信度作用不是太大，请勿过多依赖此值。
word_info	Array of Object	分词输出列表。

表 5-18 word_info 数据结构

参数名	是否必选	参数类型	说明
start_time	否	Integer	起始时间
end_time	否	Integer	结束时间
word	否	String	分词

示例

```
{
  "resp_type": "RESULT",
  "trace_id": "149edb8a-664b-4f0d-9935-847a57b1e3a5",
  "segments": [
    {
      "start_time": 0,
```


版权所有 © 华为云计算技术有限公司

```
    "end_time": 5480,
    "is_final": true,
    "result": {
      "text": "欢迎大家使用华为云语音交互服务。",
      "score": 0.8452650904655457,
      "word_info": [
        {
          "start_time": 1700,
          "end_time": 1800,
          "word": "欢"
        },
        {
          "start_time": 1940,
          "end_time": 2040,
          "word": "迎"
        },
        {
          "start_time": 2180,
          "end_time": 2280,
          "word": "大"
        },
        {
          "start_time": 2300,
          "end_time": 2400,
          "word": "家"
        },
        {
          "start_time": 2480,
          "end_time": 2580,
          "word": "使"
        },
        {
          "start_time": 2660,
          "end_time": 2760,
          "word": "用"
        },
        {
          "start_time": 2900,
          "end_time": 3000,
          "word": "华"
        },
        {
          "start_time": 3080,
          "end_time": 3180,
          "word": "为"
        },
        {
          "start_time": 3260,
          "end_time": 3360,
          "word": "云"
        },
        {
          "start_time": 3560,
          "end_time": 3660,
          "word": "语"
        },
        {
          "start_time": 3660,
          "end_time": 3720,
          "word": "音"
        },
        {
          "start_time": 3860,
          "end_time": 3960,
          "word": "交"
        },
        {
          "start_time": 4040,
          "end_time": 4140,
```


版权所有 © 华为云计算技术有限公司

```
          "word": "互"
        },
        {
          "start_time": 4280,
          "end_time": 4380,
          "word": "服"
        },
        {
          "start_time": 4460,
          "end_time": 4560,
          "word": "务"
        }
      ]
    }
  }
 ]
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.4.4 错误响应

错误响应，通常是指不影响流程，但当前会话无法再进行下去的错误，包括如下情况：

- 配置串错误，包括存在不识别的配置串，或者配置串值的范围不合法。
- 时序不正确，比如连续发送两次“开始识别”指令。
- 识别过程中发生错误，比如音频解码发生错误。

出现错误响应时，如果已经在一个会话中了，会再发送一个“结束识别”的响应，表示识别会话结束。如果会话还没有开始，那么发送此错误响应后不做其它操作。此后的音频数据都被忽略，直到收到下一个“开始识别”请求。

响应消息

表 5-19 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为ERROR，表示错误响应。
trace_id	String	服务内部的跟踪令牌，可用于在日志中追溯具体流程。 在某些错误情况下，可能没有此字段。
error_code	String	错误码列表。详细错误码解释，请参见**错误码**。
error_msg	String	返回错误信息。

版权所有 © 华为云计算技术有限公司

示例

```
{
    "resp_type": "ERROR",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "error_code": "SIS.0002",
    "error_msg": "***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

5.4.5 严重错误响应

严重错误，通常指流程无法继续的情况。比如当出现客户端分片音频间隔超时（例如20s）。

出现严重错误响应时，流程不再继续，服务器端会主动断连。

响应消息

表 5-20 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为FATAL_ERROR，表示开始识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
error_code	String	错误码列表。详细错误码解释，请参见**错误码**。
error_msg	String	返回错误信息。

示例

```
{
    "resp_type": "FATAL_ERROR",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "error_code": "SIS.0002",
    "error_msg": "***"
}
```

状态码

状态码请参见**状态码**。

版权所有 © 华为云计算技术有限公司

错误码

错误码请参见错误码。

5.4.6 结束识别请求响应

服务器端收到“结束识别”请求时或语音识别过程中发生错误，服务端会向客户端推送如下响应消息，以json字符串形式放置在text message中。

响应消息

表 5-21 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为END，表示结束识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
reason	String	结束原因，详情请参见表 结束原因表。

表 5-22 结束原因表

参数名	说明
NORMAL	正常结束。
CANCEL	用户取消，也即客户端发送“结束识别”指令时cancel参数为true。
ERROR	识别过程中发生错误。

示例

```
{
    "resp_type": "END",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "reason": "NORMAL",
}
```

状态码

状态码请参见状态码。

错误码

错误码请参见错误码。

版权所有 © 华为云计算技术有限公司

6 一句话识别

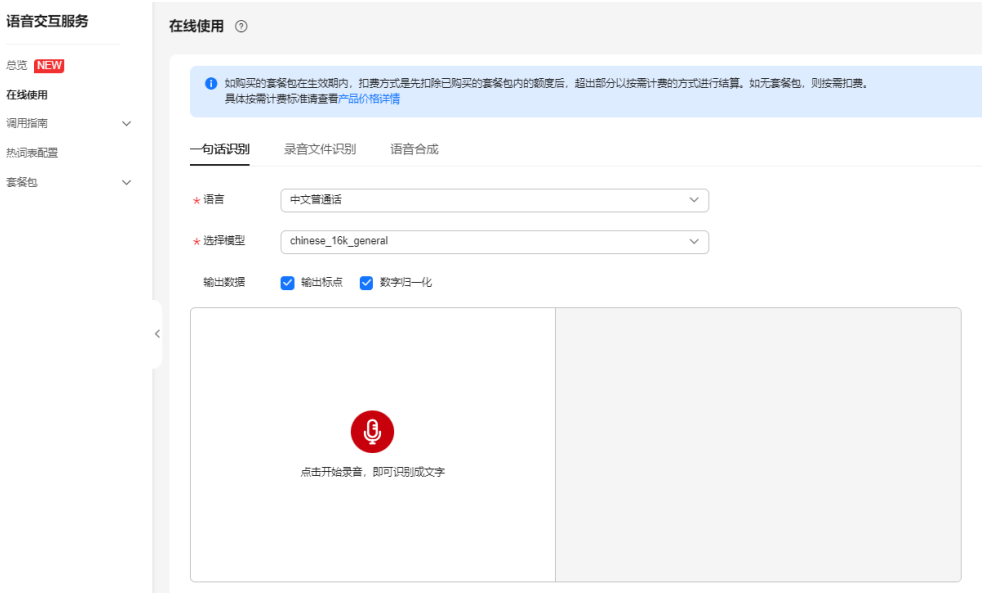
6.1 http 接口

功能介绍

一句话识别接口，用于短语音的同步识别。一次性上传1min以内音频，能快速返回识别结果。该接口的使用限制请参见**约束与限制**，详细使用指导请参见**SIS服务使用简介**章节。

SIS服务控制台提供了一句话识别的在线使用页面，用户无需编程即可使用服务功能。

图 6-1 在线使用



调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

版权所有 © 华为云计算技术有限公司

URI

POST /v1/{project_id}/asr/short-audio

表 6-1 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

请求参数

表 6-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“**企业项目管理**”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

版权所有 © 华为云计算技术有限公司

表 6-3 请求 Body 参数

参数	是否必选	参数类型	描述
config	是	**Config** object	配置信息。
data	是	String	语音数据，Base64编码，要求Base64编码后大小不超过4M，音频时长不超过1分钟。Base64编码示例如/+MgxAAUeHpMAUkQAANhuRAC...，如果携带data:audio/mp3; base64,类前缀，会产生报错。

表 6-4 Config

参数	是否必选	参数类型	描述
audio_format	是	String	支持语音的格式，请参考**表audio_format取值范围**。
property	是	String	所使用的模型特征串，通常是“语种_采样率_领域”的形式，采样率需要与音频采样率保持一致，取值范围请参考**表property取值范围**。
add_punc	否	String	表示是否在识别结果中添加标点，取值为“yes”和“no”，默认为“no”。
digit_norm	否	String	表示是否将语音中的数字识别为阿拉伯数字，取值为“yes” 和 “no”，默认为“yes”。
vocabulary_id	否	String	热词表id，不使用则不填写。创建热词表信息请参考**创建热词表**。
need_word_info	否	String	表示是否在识别结果中输出分词结果信息，取值为“yes”和“no”，默认为“no”。

表 6-5 audio_format 取值范围

audio_format取值	描述
pcm16k16bit	16k16bit单通道录音数据。
pcm8k16bit	8k16bit单通道录音数据。
ulaw16k8bit	16k8bit ulaw单通道录音数据。
ulaw8k8bit	8k8bit ulaw单通道录音数据。
alaw16k8bit	16k8bit alaw单通道录音数据。

版权所有 © 华为云计算技术有限公司

audio_format取值	描述
alaw8k8bit	8k8bit alaw单通道录音数据。
mp3	mp3格式音频。目前仅支持单通道的音频。
aac	aac格式音频。目前仅支持单通道的音频。
wav	带wav封装头的格式，从封装头中自动确定格式，目前仅支持8k/16k采样率、单通道、pcm, alaw, ulaw三种编码格式。
amr	AMR窄带(8k) 压缩录音数据。目前仅支持单通道的音频。
amrwb	AMR 宽带(16k) 压缩录音数据。目前仅支持单通道的音频。
auto	由引擎自动判断音频数据的格式并解码，支持自动判断amr，flac，m4a，mp3，ogg，webm，wav，aac，ac3，mov，wma、amrwb格式。

表 6-6 property 取值范围

property取值	描述
chinese_16k_general	支持采样率为8k/16k的中文普通话语音识别，采用新一代端到端识别算法，识别准确率更高。 区域支持cn-east-3和cn-north-4。
chinese_16k_travel	支持采样率为16k的中文普通话语音识别，对噪声场景进行了优化。若所处环境噪声较大，建议采用该参数。 区域支持cn-east-3和cn-north-4。
sichuan_16k_common	支持采样率为8k/16k的四川话方言识别，区域仅支持cn-north-4。
cantonese_16k_common	支持采样率为8k/16k的粤语方言识别，区域仅支持cn-north-4。
shanghai_16k_common	支持采样率为8k/16k的上海话方言识别，区域仅支持cn-north-4。
chinese_8k_common	支持采样率为8k的中文普通话语音识别，当明确输入音频为电话信道8k数据时，建议采用该参数。
chinese_16k_common	支持采样率为16k的中文普通话语音快速识别，采用蒸馏剪枝算法，显著提升识别速度。如果更侧重识别速度，建议采用该参数。
english_16k_common	支持采样率为8k/16k的英文语音识别，区域仅支持cn-east-3。暂不支持digit_norm参数。

版权所有 © 华为云计算技术有限公司

property取值	描述
english_8k_common	支持采样率为8k的英文语音识别，此为旧版模型，后期不再维护，建议使用english_16k_common。

响应参数

状态码： 200

表 6-7 响应 Body 参数

参数	是否必选	参数类型	描述
trace_id	是	String	服务内部的令牌，可用于在日志中追溯具体流程，调用失败无此字段。 在某些错误情况下可能没有此令牌字符串。
result	是	**Result** object	调用成功表示识别结果，调用失败时无此字段。

表 6-8 Result

参数	是否必选	参数类型	描述
text	是	String	调用成功表示识别出的内容。
score	是	Float	调用成功表示识别出的置信度，取值范围：0~1。
word_info	否	Array of **WordInfo** objects	分词信息列表。 分词是指将识别出的文本进一步拆分成单独的词语。

表 6-9 WordInfo

参数	是否必选	参数类型	描述
start_time	否	Integer	起始时间。
end_time	否	Integer	结束时间。
word	否	String	分词。

状态码： 400

版权所有 © 华为云计算技术有限公司

表 6-10 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 上传短音频，快速返回识别结果

```
POST https://{endpoint}/v1/{project_id}/asr/short-audio

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...

Request body:
{
  "config":
  {
    "audio_format": "wav",
    "property": "chinese_16k_general",
    "add_punc": "yes",
    "need_word_info": "yes"
  },
  "data": "/+MgxAAUeHpMAUkQAANhuRAC..."
}
```

- 使用Python3语言上传短音频，快速返回识别结果

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装requests，执行pip install requests
import requests
import base64
import json

def sasr_demo():
    url = 'https://{{endpoint}}/v1/{{project_id}}/asr/short-audio'  # endpoint和project_id需替换
    token = '用户对应region的token'
    file_path = '将要识别音频的路径'
    with open(file_path, 'rb') as f:
        data = f.read()
        base64_data = str(base64.b64encode(data), 'utf-8')
    header = {
        'Content-Type': 'application/json',
        'X-Auth-Token': token
    }
    body = {
        'data': base64_data,
        'config': {
            'property': 'chinese_16k_general',
            'audio_format': 'pcm16k16bit'
        }
    }
    resp = requests.post(url, data=json.dumps(body), headers=header)
    print(resp.text)

if __name__ == '__main__':
    sasr_demo()
```


版权所有 © 华为云计算技术有限公司

- 使用Java语言上传短音频，快速返回识别结果

```
import java.io.BufferedReader;
import java.io.InputStream;
import java.io.InputStreamReader;
import java.io.OutputStreamWriter;
import java.net.HttpURLConnection;
import java.net.URL;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 */
public class SasrDemo {
  public void sasrDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      URL url = new URL("https://{{endpoint}}/v1/{{project_id}}/asr/short-audio");
      String token = "对应region的token";
      String audioBody = "8k wav格式audio对应base64编码";
      HttpURLConnection connection = (HttpURLConnection)url.openConnection();
      connection.setRequestMethod("POST");
      connection.setDoInput(true);
      connection.setDoOutput(true);
      connection.addRequestProperty("Content-Type", "application/json");
      connection.addRequestProperty("X-Auth-Token", token);


      OutputStreamWriter osw = new OutputStreamWriter(connection.getOutputStream(), "UTF-8");
      String body = "{\"data\":\"" +  audioBody + "\", \"config\": { \"audio_format\": \"wav\", "
        + "\"property\":\"chinese_16k_general\"}}";
      osw.append(body);
      osw.flush();
      InputStream is = connection.getInputStream();
      BufferedReader br = new BufferedReader(new InputStreamReader(is, "UTF-8"));
      while (br.ready()) {
        System.out.println(br.readLine());
      }
    } catch (Exception e) {
      e.printStackTrace();
    }
  }
  public static void main(String[] args) {
    SasrDemo sasrDemo = new SasrDemo();
    sasrDemo.sasrDemo();
  }
}
```

响应示例

状态码：200

成功响应示例

```
{
  "trace_id": "567e8537-a89c-13c3-a882-826321939651",
  "result":{
    "text": "欢迎使用语音云服务。",
    "score": 0.9,
    "word_info": [
          {
              "start_time": 150,
              "end_time": 570,
              "word": "欢迎"
          },
          {
              "start_time": 570,
              "end_time": 990,
              "word": "使用"
          },
```

```
            {
                "start_time": 990,
                "end_time": 1380,
                "word": "语音"
            },
            {
                "start_time": 1380,
                "end_time": 1590,
                "word": "云"
            },
            {
                "start_time": 1590,
                "end_time": 2070,
                "word": "服务"
            }
        ]
    }
}
```

状态码：400

失败响应示例

```
{
    "error_code":"SIS.0001",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

6.2 websocket 接口

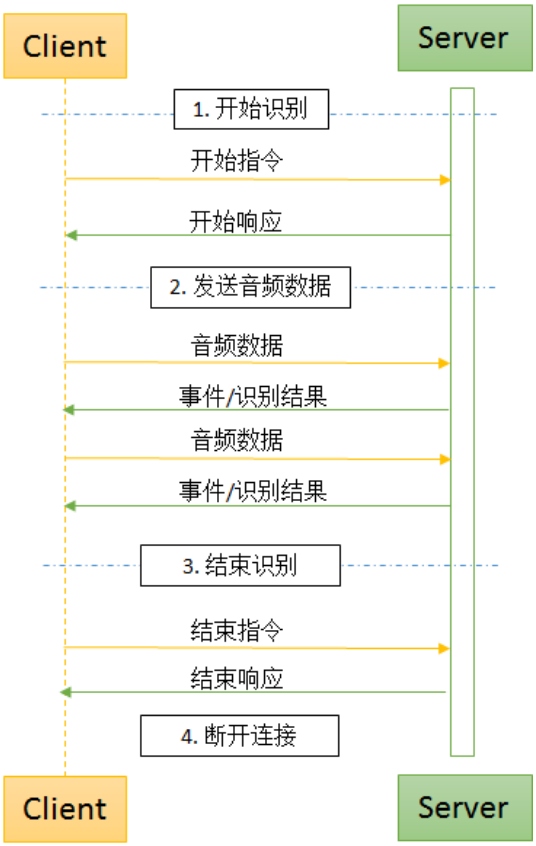
功能介绍

一句话识别websocket接口支持识别1min以内的音频，交互过程如**图 客户端和服务端交互流程**所示，主要分为开始识别、发送音频数据，结束识别、断开连接四个步骤。

websocket接口同http接口一致按次计费，只要建立连接成功，发送音频，服务开始识别，则本次调用计费生效。如果用户发送错误end请求或者持续20s未发送音频而产生了报错，该次调用依然认为生效。如果连接成功后未发送音频直接断开，或者请求字段不正确而产生异常，则认为本次调用无效，不会纳入计费次数。

版权所有 © 华为云计算技术有限公司

图 6-2 客户端和服务端交互流程



wss-URI

- wss-URI格式

  wss /v1/{project_id}/asr/short-audio

- 参数说明

表 6-11 参数说明

参数名	是否必选	说明
project_id	是	项目编号。获取方法，请参见**获取项目ID**。

表 6-12 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“企业项目管理”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

开始识别

- 功能介绍

  当wss握手请求收到成功响应后，客户端到服务端的通信协议会升级为Websocket协议。通过Websocket协议，客户端发送开始识别请求，用于配置一句话识别的配置信息。

- 请求消息

表 6-13 参数说明

参数名	是否必选	参数类型	说明
command	是	String	表示客户端发送开始识别请求，参数值需设置为START。
config	是	Object	配置信息。结构信息请参见表config数据结构。

版权所有 © 华为云计算技术有限公司

表 6-14 config 数据结构

参数	是否必选	参数类型	说明
audio_format	是	String	支持语音的格式，请参见**表audio_format取值范围**。
property	是	String	所使用的模型特征串。通常是“语种_采样率_领域”的形式，例如chinese_8k_common。请参见**表property取值范围**。
add_punc	否	String	表示是否在识别结果中添加标点，取值为yes 、 no，默认no。
digit_norm	否	String	表示是否将语音中的数字识别为阿拉伯数字，取值为yes 、 no，默认为yes。识别结束后，会将数字识别为阿拉伯数字。
interim_results	否	String	是否输出中间结果，可以为yes或no。默认为no，表示不输出中间结果。
vocabulary_id	否	String	热词表id，不使用热词则不填写。 创建热词表信息请参考**创建热词表**。
need_word_info	否	String	表示是否在识别结果中输出分词结果信息，取值为“yes”和“no”，默认为“no”。

表 6-15 property 取值范围

property取值	说明
chinese_8k_general	支持采样率为8k的中文普通话语音识别，采用新一代端到端识别算法，识别准确率更高。 格式支持pcm8k16bit/alaw8k8bit/ulaw8k8bit，区域支持cn-east-3和cn-north-4（强烈推荐使用）。
chinese_16k_general	支持采样率为16k的中文普通话语音识别，采用新一代端到端识别算法，识别准确率更高。 格式支持pcm16k16bit/alaw16k8bit/ulaw16k8bit，区域支持cn-east-3和cn-north-4（强烈推荐使用）。
chinese_8k_common	支持采样率为8k的中文普通话语音识别，此为旧版引擎，建议使用chinese_8k_general。

版权所有 © 华为云计算技术有限公司

property取值	说明
chinese_16k_common	支持采样率为16k的中文普通话语音识别，此为旧版引擎，建议使用chinese_8k_general。
sichuan_16k_common	支持采样率为16k的中文普通话与四川话方言语音识别。区域仅支持cn-north-4。
cantonese_16k_common	支持采样率为16k的粤语方言语音识别。区域仅支持cn-north-4。
shanghai_16k_common	支持采样率为16k的上海话方言语音识别。区域仅支持cn-north-4。

表 6-16 audio_format 取值范围

audio_format取值	说明
pcm16k16bit	16k16bit单通道录音数据。
pcm8k16bit	8k16bit单通道录音数据。
ulaw16k8bit	16k8bit ulaw单通道录音数据。
ulaw8k8bit	8k8bit ulaw单通道录音数据。
alaw16k8bit	16k8bit alaw单通道录音数据。
alaw8k8bit	8k8bit alaw单通道录音数据。

说明

目前仅支持裸音频格式，仅支持pcm编码的wav格式，不支其他wav头或者arm格式的编码。

- 示例

```
{
  "command": "START",
  "config":
  {
    "audio_format": "pcm8k16bit",
    "property": "chinese_8k_general",
    "add_punc": "yes",
    "interim_results": "yes",
    "need_word_info": "yes"
  }
}
```

- 状态码

  状态码请参见**状态码**。

- 错误码

  错误码请参见**错误码**。

版权所有 © 华为云计算技术有限公司

发送音频数据

在收到“开始识别”的响应之后，可以开始发送音频数据。为节省流量，音频以二进制数据帧形式（binary message）的方式发送。

音频数据将分片发送，也即在获得一定量音频数据的同时就可以发送一个binary message，每个分片建议在50ms~1000ms之间，建议在需要实时反馈的情况下100ms，不需要实时反馈的情况下500ms。

结束识别

- 功能介绍

  对于识别中的对话，需要在Websocket上发送“结束识别”的请求来取消或结束识别。 "结束识别"请求使用文本类型的数据帧(text message)发送，命令和参数以json字符串的形式提供。

- 请求消息

  表 6-17 参数说明

参数名	是否必选	参数类型	说明
command	是	String	表示客户端结束识别请求，参数值设置为END。

- 示例

```
{
  "command": "END"
}
```

- 状态码

  状态码请参见**状态码**。

- 错误码

错误码请参见**错误码**。

响应结果

- 开始识别响应

  由于WebSocket是全双工的，因此响应就是从服务器端发送给客户端的消息，但也并不是所有的请求信息都有一条对应的响应。服务器端收到“开始识别”请求时，会给出如下响应消息，以json字符串形式放置在text message中。

  表 6-18 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为START，表示开始识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。

版权所有 © 华为云计算技术有限公司

示例

```
{
    "resp_type": "START",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651"
}
```

- 结果响应

  服务端在收到客户端发送的连续音频数据后， 当服务端识别出结果后会实时向客户端按句推送识别结果响应消息， 以json字符串形式放置在text message中。

表 6-19 响应参数

参数名	参数类型	说明
resp_type	String	响应类型。参数值为RESULT，表示识别结果响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
segments	Array of objects	多句结果。 请参考**表 segment 数据结构**。

表 6-20 segment 数据结构

参数名	参数类型	说明
start_time	Integer	一句的起始时间戳，单位为ms。
end_time	Integer	一句的结束时间戳，单位为ms。
is_final	Boolen	true表示是最终结果， false表示为中间临时结果。
result	Object	调用成功表示识别结果，调用失败时无此字段。 请参考**表 result数据结构**。

表 6-21 result 数据结构

参数名	参数类型	说明
text	String	识别结果。

版权所有 © 华为云计算技术有限公司

参数名	参数类型	说明
score	Float	识别结果的置信度，取值范围：0~1。由于中间结果会不断刷新和纠正，此置信度会不断变化，直到返回当前句子的最终结果，置信度才固定不变。 **说明** 目前置信度作用不是太大，请勿过多依赖此值。
word_info	Array of Object	分词输出列表。

表 6-22 word_info 数据结构

参数名	是否必选	参数类型	说明
start_time	否	Integer	起始时间
end_time	否	Integer	结束时间
word	否	String	分词

示例

```
{
  "resp_type": "RESULT",
  "trace_id": "149edb8a-664b-4f0d-9935-847a57b1e3a5",
  "segments": [
    {
      "start_time": 0,
      "end_time": 5480,
      "is_final": true,
      "result": {
        "text": "欢迎大家使用华为云语音交互服务。",
        "score": 0.8452650904655457,
        "word_info": [
          {
            "start_time": 1700,
            "end_time": 1800,
            "word": "欢"
          },
          {
            "start_time": 1940,
            "end_time": 2040,
            "word": "迎"
          },
          {
            "start_time": 2180,
            "end_time": 2280,
            "word": "大"
          },
          {
            "start_time": 2300,
            "end_time": 2400,
            "word": "家"
          },
          {
```


版权所有 © 华为云计算技术有限公司

```
          "start_time": 2480,
          "end_time": 2580,
          "word": "使"
        },
        {
          "start_time": 2660,
          "end_time": 2760,
          "word": "用"
        },
        {
          "start_time": 2900,
          "end_time": 3000,
          "word": "华"
        },
        {
          "start_time": 3080,
          "end_time": 3180,
          "word": "为"
        },
        {
          "start_time": 3260,
          "end_time": 3360,
          "word": "云"
        },
        {
          "start_time": 3560,
          "end_time": 3660,
          "word": "语"
        },
        {
          "start_time": 3660,
          "end_time": 3720,
          "word": "音"
        },
        {
          "start_time": 3860,
          "end_time": 3960,
          "word": "交"
        },
        {
          "start_time": 4040,
          "end_time": 4140,
          "word": "互"
        },
        {
          "start_time": 4280,
          "end_time": 4380,
          "word": "服"
        },
        {
          "start_time": 4460,
          "end_time": 4560,
          "word": "务"
        }
      ]
    }
  }
 ]
}
```


- 错误响应

  错误响应，包括如下情况：

  – 配置串错误，包括存在不识别的配置串，或者配置串值的范围不合法。
  – 时序不正确，比如连续发送两次“开始识别”指令。
  – 识别过程中发生错误，比如音频解码发生错误。

表 6-23 响应参数

参数名	参数类型	说明
resp_type	String	参数值为ERROR，表示错误响应。
trace_id	String	服务内部的跟踪令牌，可用于在日志中追溯具体流程。 在某些错误情况下，可能没有此字段。
error_code	String	错误码列表。详细错误码解释，请参见**错误码**。
error_msg	String	返回错误信息。

示例

```
{
    "resp_type": "ERROR",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "error_code": "SIS.0002",
    "error_msg": "***"
}
```

- 结束识别响应

  服务器端收到“结束识别”请求时或语音识别过程中发生错误，服务端会向客户端推送如下响应消息，以json字符串形式放置在text message中。

表 6-24 响应参数

参数名	参数类型	说明
resp_type	String	参数值为END，表示结束识别响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
reason	String	结束原因，详情请参见**表 结束原因表**。

表 6-25 结束原因表

参数名	说明
NORMAL	正常结束。
ERROR	识别过程中发生错误。

示例

```
{
    "resp_type": "END",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "reason": "NORMAL",
}
```


版权所有 © 华为云计算技术有限公司

代码示例

```
import okhttp3.OkHttpClient;
import okhttp3.Request;
import okhttp3.Response;
import okhttp3.WebSocket;
import okhttp3.WebSocketListener;
import okio.ByteString;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 * 使用前需已配置okhttp、okio jar包。jar包可通过下载SDK获取。
 */
public class SasrWebsocketDemo {
  public void sasrWebsocketDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      String url = "wss://{{endpoint}}/v1/{{project_id}}/asr/short-audio";
      String token = "对应region的token";
      byte[] data = null;  // 存放将要发送音频的byte数组
      OkHttpClient okHttpClient = new OkHttpClient();
      Request request = new Request.Builder().url(url).header("X-Auth-Token", token).build();
      WebSocket webSocket = okHttpClient.newWebSocket(request, new MyListener());
      webSocket.send("{\"command\": \"START\", \"config\": {\"audio_format\": \"pcm8k16bit\", \"property\": 
\"chinese_8k_general\"}}");
      webSocket.send(ByteString.of(data));  // audio太大注意要分片发送，否则会报错。建议分片大小3200
      webSocket.send("{  \"command\": \"END\"}");
      Thread.sleep(10000);
      webSocket.close(1000, null);

    } catch (Exception e) {
      e.printStackTrace();
    }

  }

  class MyListener extends WebSocketListener {
    @Override

    public void onOpen(WebSocket webSocket, Response response) {
      System.out.println("conneected");
    }

    @Override
    public void onClosed(WebSocket webSocket, int code, String reason) {

      System.out.println("closed");
    }

    @Override
    public void onFailure(WebSocket webSocket, Throwable t, Response response) {
      t.printStackTrace();
    }

    @Override
    public void onMessage(WebSocket webSocket, String text) {
      System.out.println(text);
    }

  }

  public static void main(String[] args) {
    SasrWebsocketDemo sasrWebsocketDemo = new SasrWebsocketDemo();
    sasrWebsocketDemo.sasrWebsocketDemo();
  }
}
```

7 录音文件识别接口

7.1 提交录音文件识别任务

录音文件识别接口，用于识别长录音文件，支持中文普通话、四川话识别。录音文件放在OBS（对象存储服务）或公网可访问的服务器上。

由于录音文件识别通常会需要较长的时间，因此识别是异步的，即接口分为创建识别任务和查询任务状态两个接口。创建识别任务接口创建任务完成后返回job_id参数，然后用户通过调用**录音文件识别状态查询**接口来获得转写状态和结果。

功能介绍

该接口用于提交录音文件识别任务，其中录音文件保存在用户的OBS桶中或公网可访问的服务器上（需保证可使用域名访问）。用户开通录音识别服务时，如果录音文件存放在OBS桶中，需授权录音文件引擎读取用户OBS桶权限，授权方法见**配置OBS访问权限**。该接口的使用限制请参见**约束与限制**，详细使用指导请参见**SIS服务使用简介**章节。

SIS服务控制台提供了录音文字识别的在线使用页面，用户无需编程即可使用服务功能。

图 7-1 在线使用

版权所有 © 华为云计算技术有限公司

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

接口约束

- 录音时长不超过5小时，文件大小不超过300M。且音频时长至少0.01s以上。
- 需要将识别的录音文件存放在某服务上，可以通过URL访问。
  - 推荐使用华为云OBS：授权配置请参见**配置OBS服务**，优先推荐使用与请求服务同一region的私有桶文件，如果region不一致，可考虑使用**obs签名**。
  - 您也可以把录音文件放在自行搭建服务器上，提供下载文件的地址。URL不能使用IP地址，只能使用域名，请尽量避免中文。
- 音频转写时长受音频时长和排队任务数量影响，音频时长和理论返回时间可参见**表 音频转写时长参考表**。如果转写耗时比理论时延长，大概率表示当前时间段出现转写高峰，请耐心等待，我们承诺最大转写时长不超过6小时。

表 7-1 音频转写时长参考表

音频时长	参考返回时间
小于10分钟	小于2分钟
10分钟-60分钟	2分钟-12分钟
大于60分钟	12 分钟-60分钟

- 识别结果保存72小时（从识别完成的时间算起）。72小时后如果再访问，将会返回 "task id is not found"错误。

URI

POST /v1/{project_id}/asr/transcriber/jobs

表 7-2 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

版权所有 © 华为云计算技术有限公司

请求参数

表 7-3 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“**企业项目管理**”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

表 7-4 请求 Body 参数

参数	是否必选	参数类型	描述
config	是	**Transcriber Config** object	录音文件识别配置信息。
data_url	是	String	存放录音文件地址： ● 推荐使用华为云OBS：授权配置请参见**配置OBS服务**。 ● 您也可以把录音文件放在自行搭建服务器上，提供下载文件的地址。URL不能使用IP地址，只能使用域名，请尽量避免中文。

版权所有 © 华为云计算技术有限公司

表 7-5 TranscriberConfig

参数	是否必选	参数类型	描述
audio_format	否	String	支持语音的格式，请参考**表audioformat取值范围**。
property	是	String	所使用的模型特征串。通常是“语种_采样率_领域”的形式，例如chinese_8k_common。 采样率需要与音频采样率保持一致。 当前支持如下模型特征串： ● chinese_8k_general（电话通话领域，区域支持cn-north-4, cn-east-3，最新端到端通用模型，强烈推荐使用） ● chinese_16k_media（音视频领域，区域支持cn-north-4，cn-east-3，强烈推荐使用） ● chinese_8k_common（中文普通话语音识别，区域支持cn-north-4, cn-east-3） ● chinese_16k_conversation（会议场景的中文普通话语音识别，区域支持cn-north-4, cn-east-3） ● sichuan_8k_common（四川话识别，区域支持cn-north-4，cn-east-3） ● shanghai_8k_common（上海话识别，区域支持cn-north-4, cn-east-3） ● cantonese_8k_common（粤语识别，区域支持cn-north-4, cn-east-3） ● chinese_8k_bank（银行领域，区域仅支持cn-north-4） ● chinese_8k_insurance（保险领域，区域仅支持cn-north-4）
add_punc	否	String	表示是否在识别结果中添加标点，取值为yes 、 no，默认no。
callback_url	否	String	表示回调 url，即用户用于接收识别结果的服务器地址，不支持IP地址方式调用，url长度小于2048字节。服务请求方法为POST，请求体参见**表 响应参数**。 ● 如果用户使用回调方式获取识别结果，需填写该参数，处理成功后用户服务器需返回状态码“200”。 ● 如果用户使用轮询方式获取识别结果，则无需填写该参数。
digit_norm	否	String	表示是否将语音中的数字识别为阿拉伯数字，取值为yes 、 no，默认为yes。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
need_analysis_info	否	**AnalysisInfo** object	该参数用于设置质检分析参数，是否需要输出质检分析内容以及会议场景多说话人分离，如果为null，表示不需要。
vocabulary_id	否	String	热词表id，不使用则不填写。 创建热词表信息请参考**创建热词表**。
need_word_info	否	String	表示是否在识别结果中输出分词结果信息，取值为“yes”和“no”，默认为“no”。

表 7-6 AnalysisInfo

参数	是否必选	参数类型	描述
diarization	否	Boolean	是否需要做话者分离。 默认为true，表示会进行话者分离，识别结果中会包含role项（角色）。 如果diarization为false，那么结果中不会出现role项。 **说明** 本表中的参数对audio_format格式中采样率为8k和16k的音频有效。对于8k单声道音频，8k双声道音频以及16k双声道音频仅支持两个说话人分离；对于16k单声道音频，自动识别说话人个数，返回结果最多支持10个说话人，并且property仅支持chinese_16k_media。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
channel	否	String	语音文件声道信息，默认为MONO，可设置为MONO，LEFT_AGENT, RIGHT_AGENT。 如果channel 为MONO，那么原始文件需要为单声道文件；如果channel 为MONO，diarization为true时，会默认输出speed结果。 如果为双声道文件，系统会将其转换成单声道文件，可能会影响识别效果。 如果 channel 为 LEFT_AGENT或RIGHT_AGENT, 则原始文件需要为双声道文件，如果为单声道文件，系统会将其转换成双声道文件，识别结果会出现两条内容完全一致的文本。 当channel 为 LEFT_AGENT或RIGHT_AGENT，且diarization为true时，系统会按照配置给出对应角色。其中： ● LEFT_AGENT 指定左声道语音为agent（坐席）。 ● RIGHT_AGENT 指定右声道为agent（坐席）。
emotion	否	Boolean	是否需要做情绪检测，默认为true。仅对audio_format格式中采样率为8k的音频有效。 目前支持NORMAL（正常）、ANGRY（愤怒）和UNKNOWN（未知）情绪识别。
speed	否	Boolean	是否需要输出语速信息，默认为true。

表 7-7 audio_format 取值范围

audio_format取值	描述
auto	默认格式，系统自动判断，支持的音视频格式如下：WAV（内部支持pcm/ulaw/alaw/adpcm编码格式）、AMR、FLAC、M4A、MP3、OGG、WEBM、AAC、AC3、MOV、WMA、MP4、AVI、RMVB、MKV、FLV、F4V、WMV、3GP。支持双声道的音视频。
pcm16k16bit	16k16bit裸音频录音数据。
pcm8k16bit	8k16bit裸音频录音数据。
ulaw16k8bit	16k8bit ulaw 裸音频录音数据。

版权所有 © 华为云计算技术有限公司

audio_format取值	描述
ulaw8k8bit	8k8bit ulaw 裸音频录音数据。
alaw16k8bit	16k8bit alaw 裸音频录音数据。
alaw8k8bit	8k8bit alaw 裸音频录音数据。

响应参数

状态码： 200

表 7-8 响应 Body 参数

参数	是否必选	参数类型	描述
job_id	否	String	创建的任务标识, 如果创建任务成功时必须存在。 由于录音文件识别通常会需要较长的时间，因此识别是异步的，即接口分为创建识别任务和查询任务状态两个接口。 创建识别任务接口创建任务完成后返回job_id参数，然后用户通过调用**录音文件识别状态查询**接口来获得转写状态和结果。

状态码： 400

表 7-9 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 提交录音文件识别任务

```
POST https://{endpoint}/v1/1609d8170cd86660a81281e5a3a03f8b/asr/transcriber/jobs

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```


版权所有 © 华为云计算技术有限公司

```
Request Body:
{
  "config":
  {
    "audio_format": "auto",
    "property": "chinese_8k_general",
    "add_punc": "yes",
    "need_analysis_info":
    {
      "diarization": true,
      "channel": "LEFT_AGENT",
      "emotion": true
    },
    "need_word_info": "yes"
  },
  "data_url":"https://****_voip.obs.myxxxcloud.com/VOIP/***.wav"
}
```

- 使用Python3语言提交录音文字识别任务

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装requests，执行pip install requests
import requests
import time
import json

def lasr_demo():
    url = 'https://{{endpoint}}/v1/{{project_id}}/asr/transcriber/jobs'  # endpoint和project_id需替换
    token = '用户对应region的token'
    obs_url = 'obs路径'
    header = {
        'Content-Type': 'application/json',
        'X-Auth-Token': token
    }
    body = {
        'data_url': obs_url,
        'config': {
            'property': 'chinese_8k_general',
            'audio_format': 'auto'
        }
    }
    resp = requests.post(url, data=json.dumps(body), headers=header)
    job_id = json.loads(resp.text, encoding='utf-8')['job_id']
    query_task(header, job_id)
def query_task(header, job_id):
    # endpoint和project_id需替换
    url = 'https://{{endpoint}}/v1/{{project_id}}/asr/transcriber/jobs/' + job_id
    for _ in range(1000):
        resp = requests.get(url, headers=header)
        print(resp.text)
        if resp.text.find('segments') != -1:
            break
        time.sleep(1)
if __name__ == '__main__':
    lasr_demo()
```

- 使用Java语言提交录音文件识别任务

```
import java.io.BufferedReader;
import java.io.InputStream;
import java.io.InputStreamReader;
import java.io.OutputStreamWriter;
import java.net.HttpURLConnection;
import java.net.URL;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 */
public class LasrDemo {
  public void lasrDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
```


版权所有 © 华为云计算技术有限公司

```
    URL url = new URL("https://{{endpoint}}/v1/{{project_id}}/asr/transcriber/jobs");
    String token = "对应region的token";
    String obsUrl = "音频在obs的路径";
    HttpURLConnection connection = (HttpURLConnection)url.openConnection();
    connection.setRequestMethod("POST");
    connection.setDoInput(true);
    connection.setDoOutput(true);
    connection.addRequestProperty("Content-Type", "application/json");
    connection.addRequestProperty("X-Auth-Token", token);

    OutputStreamWriter osw = new OutputStreamWriter(connection.getOutputStream(), "UTF-8");
    String body = "{\"data_url\":\"" +  obsUrl + "\", \"config\": { \"audio_format\": \"auto\", "
      + "\"property\":\"chinese_8k_general\"}}";
    osw.append(body);
    osw.flush();
    InputStream is = connection.getInputStream();
    BufferedReader br = new BufferedReader(new InputStreamReader(is, "UTF-8"));
    String jobId = "";
    while (br.ready()) {
      String line = br.readLine().trim();
      jobId = line.replace("{\"job_id\":\"", "").replace("\"}", "");
    }
    queryTask(token, jobId);
  } catch (Exception e) {
    e.printStackTrace();
  }
}
private void queryTask(String token, String jobId) {
  try {
    // endpoint和projectId需要替换成实际信息。
    URL url = new URL("https://{{endpoint}}/v1/{{project_id}}/asr/transcriber/jobs/" + jobId);
    while (true) {
      HttpURLConnection connection = (HttpURLConnection)url.openConnection();
      connection.setRequestMethod("GET");
      connection.setDoInput(true);
      connection.setDoOutput(true);
      connection.addRequestProperty("Content-Type", "application/json");
      connection.addRequestProperty("X-Auth-Token", token);

      InputStream is = connection.getInputStream();
      BufferedReader br = new BufferedReader(new InputStreamReader(is, "UTF-8"));
      if (br.ready()) {
        String line = br.readLine();
        System.out.println(line);
        if (line.contains("segments")) {
          break;
        }
      }
      Thread.sleep(2000);
    }
  } catch (Exception e) {
    e.printStackTrace();
  }
}
public static void main(String[] args) {
  LasrDemo lasrDemo = new LasrDemo();
  lasrDemo.lasrDemo();
}
}
```

响应示例

状态码：200

成功响应示例

```
{
"job_id": "567e6536-a89c-13c3-a882-826321939656"
}
```


版权所有 © 华为云计算技术有限公司

状态码：400

失败响应示例

```
{
"error_code":"SIS.0002",
"error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

7.2 获取录音文件识别结果

功能介绍

该接口用于获取录音文件识别结果及识别状态。该接口的使用限制请参见**约束与限制**，详细使用指导请参见**SIS服务使用简介**章节。

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

GET /v1/{project_id}/asr/transcriber/jobs/{job_id}

表 7-10 路径参数

参数	是否必选	参数类型	描述
job_id	是	String	录音文件识别任务标识符。
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

版权所有 © 华为云计算技术有限公司

请求参数

表 7-11 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

响应参数

状态码： 200

表 7-12 响应 Body 参数

参数	是否必选	参数类型	描述
job_id	否	String	录音文件识别任务标识符。 ● 使用“callback_url”回调url时，该字段会随结果发送至用户服务器。 ● 使用get接口查询，不会出现该字段。
status	否	String	当前识别状态。具体状态如下所示： WAITING 等待识别。 FINISHED 识别已经完成。 ERROR 识别过程中发生错误。
create_time	否	String	任务创建时间，遵循 RFC 3339格式。 格式示例：2018-12-04T13:10:29.310Z。
start_time	否	String	开始识别时间，遵循 RFC 3339格式。 当status为FINISHED或ERROR时存在。 格式示例：2018-12-04T13:10:29.310Z。
finish_time	否	String	识别完成时间，遵循 RFC 3339格式。 当status为FINISHED或ERROR时存在。 格式示例：2018-12-04T13:10:29.310Z。
audio_duration	否	Integer	提交音频的时长，单位ms。
segments	否	Array of **Segment** objects	识别结果, 多句结果的数组。

版权所有 © 华为云计算技术有限公司

表 7-13 Segment

参数	是否必选	参数类型	描述
start_time	是	Integer	一句的起始时间戳，单位ms。
end_time	是	Integer	一句的结束时间戳，单位ms。
result	是	**TranscriberResult** object	调用成功表示识别结果，调用失败时无此字段。数据结构参见**result**表。

表 7-14 TranscriberResult

参数	是否必选	参数类型	描述
text	是	String	识别结果文本。
analysis_info	否	**AnalysisInfoResult** object	每一句的质检分析结果对象。 仅在识别配置中的need_analysis_info不为null时存在该返回结果。数据结构参见analysisinfo数据表。
word_info	否	Array of **WordInfo** objects	分词输出列表。

表 7-15 AnalysisInfoResult

参数	是否必选	参数类型	描述
role	否	String	角色类型，目前对于8k音频以及16k双声道音频仅支持 AGENT（座席），USER（用户）。可用于电话质检等场景。对于16k单声道音频，支持多个说话人，取值范围是（speaker1 ~ speaker10），最多支持10个说话人，可用于会议场景。
emotion	否	String	情绪类型，目前仅支持NORMAL（正常），ANGRY（愤怒），UNKNOWN（未知）。 在识别配置中emotion为true时存在。
speed	否	Float	语速信息，单位是每秒字数。 在识别配置中speed为true时存在。

版权所有 © 华为云计算技术有限公司

表 7-16 WordInfo

参数	是否必选	参数类型	描述
start_time	否	Integer	起始时间
end_time	否	Integer	结束时间
word	否	String	分词

状态码： 400

表 7-17 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 获取录音文件识别结果

```
GET https://{endpoint}/v1/1609d8170cd86660a81281e5a3a03f8b/asr/transcriber/jobs/567e6536-a89c-13c3-a882-826321939656

Request header:
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

响应示例

状态码：200

成功响应示例

- 等待识别的响应示例

```
{
"status": "WAITING",
"create_time": "2018-12-04T13:10:29.310Z"
}
```

- 识别完成的响应示例

```
{
  "status": "FINISHED",
  "create_time": "2018-12-04T13:10:29.310Z",
  "start_time": "2018-12-04T13:16:29.310Z",
  "finish_time": "2018-12-04T13:35:18.310Z",
  "audio_duration":7000,
  "segments":
  [
    {
      "start_time": 100,
```


版权所有 © 华为云计算技术有限公司

```
      "end_time": 1500,
      "result":
      {
        "text": "第一句结果。",
        "word_info": [
                    {
                        "start_time": 100,
                        "word": "第一"
                    },
                    {
                        "start_time": 800,
                        "end_time": 1000,
                        "word": "句"
                    },
                    {
                        "start_time": 1000,
                         "end_time": 1500,
                         "word": "结果"
                    }
         ],
        "analysis_info":
        {
          "role": "AGENT",
          "emotion": "NORMAL",
          "speed": 5.62
        }
      }
    }
  ]
}
```

状态码：400

失败响应示例

```
{
"error_code": "SIS.30006",
"error_msg": "Invalid data found when processing input",
"status": "ERROR",
"create_time": "2019-05-08 T11:01:41Z",
"start_time": "2019-06-06T15:10:12.423Z",
"finish_time": "2019-06-06T15:10:12.424Z"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

版权所有 © 华为云计算技术有限公司

8 录音文件识别极速版接口

功能介绍

录音文件识别极速版接口，用于录音文件的同步识别。音频需要上传华为云OBS，提供华为云OBS对象地址，能快速返回识别结果。该接口的使用限制请参见**约束与限制**，详细使用指导请参见**SIS服务使用简介**章节。

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

POST /v1/{project_id}/asr/flash

表 8-1 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

表 8-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“企业项目管理”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

表 8-3 Query 参数

参数	是否必选	参数类型	描述
audio_format	是	String	支持语音的格式，请参考**表audio_format取值范围**。
property	是	String	所使用的模型特征串，通常是“语种_采样率_领域”的形式，采样率需要与音频采样率保持一致，取值范围请参考**表property取值范围**。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
obs_bucket_name	是	String	表示在OBS对象桶名，使用前请先授权配置请参见**配置OBS访问权限**。obs_bucket_name字符长度区间为[3,64)，需要进行urlencode编码。 OBS桶区域需要与接口的地址区域一致。 示例： obs url为https://sis-audio-cn4.obs.cn-north-4.myhuaweicloud.com/abc/16k_pcm.wav 则obs_bucket_name值为sis-audio-cn4，obs_object_key值为abc/16k_pcm.wav。如果上传至OBS桶中的文件或OBS路径为中文，obs_object_key的值需要采用中文形式。
obs_object_key	是	String	表示OBS对象桶中的对象的键值，长度小于1024个字符，需要进行urlencode编码。 OBS桶区域需要与接口的地址区域一致。 示例： obs url为https://sis-audio-cn4.obs.cn-north-4.myhuaweicloud.com/abc/16k_pcm.wav 则obs_bucket_name=sis-audio-cn4，obs_object_key=abc/16k_pcm.wav
add_punc	否	String	表示是否在识别结果中添加标点，取值为“yes”和“no”，默认为“no”。
digit_norm	否	String	表示是否将语音中的数字识别为阿拉伯数字，取值为“yes” 和 “no”，默认为“yes”。
need_word_info	否	String	表示是否在识别结果中输出分词结果信息，取值为“yes”和“no”，默认为“no”。
first_channel_only	否	String	表示是否在识别中只识别首个声道的音频数据，取值为“yes”和“no”，默认为“no”。 **说明** 如果音频数据为多声道，参数取默认值no时，会叠加计费。例如，双声道收费为双倍。

版权所有 © 华为云计算技术有限公司

表 8-4 audio_format 取值范围

audio_format取值	描述
wav	wav格式音频。
mp3	mp3格式音频。
m4a	m4a格式音频。
acc	acc格式音频。
opus	opus格式音频。

表 8-5 property 取值范围

property取值	描述
chinese_8k_common	支持采样率为8k的中文普通话语音识别。
chinese_16k_conversation	支持采样率为16k的会议场景的中文普通话语音识别。

请求参数

表 8-6 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

响应参数

状态码： 200

表 8-7 响应参数

参数	是否必选	参数类型	描述
trace_id	是	String	trace_id可用于在日志中追溯具体流程，调用失败无此字段。 在某些错误情况下可能没有此令牌字符串。
audio_duration	是	Integer	音频时长，单位毫秒。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
flash_result	是	Array of **FlashResult** objects	调用成功表示识别结果，调用失败时无此字段。

表 8-8 FlashResult

参数	是否必选	参数类型	描述
channel_id	否	Integer	声道Id。
sentences	否	Array of **Sentences** objects	分句信息列表。

表 8-9 Sentences

参数	是否必选	参数类型	描述
start_time	否	Integer	一句话开始时间，单位毫秒。
result	否	**Result** object	分句结果信息。
end_time	否	Integer	一句话结束时间，单位毫秒。

表 8-10 Result

参数	是否必选	参数类型	描述
text	是	String	调用成功表示识别出的内容。
score	是	Double	调用成功表示识别出的置信度（0-1之间）。
word_info	否	Array of **WordInfo** objects	分词信息列表。

表 8-11 WordInfo

参数	是否必选	参数类型	描述
start_time	否	Integer	起始时间。
end_time	否	Integer	结束时间。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
word	否	String	分词。

状态码： 400

表 8-12 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 快速识别录音文件的音频内容，转写为文字

```
POST https://{endpoint}/v1/{project_id}/asr/flash?
property=chinese_8k_common&audio_format=wav&add_punc=yes&digit_norm=yes&need_word_info=yes&first_channel_only=yes&vocabulary_id=4c111ee8-fbd6-4a73-941a-9f869c6c56f5&obs_bucket_name=test&obs_object_key=data/0601/test.wav
Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

响应示例

状态码：200

成功响应示例

```
{
    "trace_id": "567e8537-a89c-13c3-a882-826321939xxx",
    "audio_duration": 1010,
    "flash_result": [
        {
            "channel_id": 0,
            "sentences": [
                {
                    "start_time": 550,
                    "end_time": 900,
                    "result": {
                        "text": "北京",
                        "score": 0.831,
                        "word_info": [
                            {
                                "word": "北",
                                "start_time": 560,
                                "end_time": 720
                            },
                            {
                                "word": "京",
                                "start_time": 730,
```


版权所有 © 华为云计算技术有限公司

```
                    "end_time": 890
                }
            ]
        }
    }
  ]
}
```

状态码：400

失败响应示例

```
{
    "error_code":"SIS.0001",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

版权所有 © 华为云计算技术有限公司

9 语音合成接口

9.1 语音合成

功能介绍

语音合成，是一种将文本转换成逼真语音的服务。用户通过实时访问和调用API获取语音合成结果，将用户输入的文字合成为音频。通过音色选择、自定义音量、语速，为企业和个人提供个性化的发音服务。该接口的使用限制请参见**约束与限制**，详细使用指导请参见**SIS服务使用简介**章节。

SIS服务控制台提供了语音合成的在线使用页面，用户无需编程即可使用服务功能。

图 9-1 在线使用

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

版权所有 © 华为云计算技术有限公司

URI

POST /v1/{project_id}/tts

表 9-1 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

请求参数

表 9-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“**企业项目管理**”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

版权所有 © 华为云计算技术有限公司

表 9-3 请求 Body 参数

参数	是否必选	参数类型	描述
text	是	String	待合成的文本，文本长度限制不大于500字符。
config	否	**TtsConfig** object	语音合成配置信息。

表 9-4 TtsConfig

参数	是否必选	参数类型	描述
audio_format	否	String	语音格式头：wav、mp3、pcm 默认：wav 父节点：config
sample_rate	否	String	采样率：16000、8000赫兹 默认：8000 父节点：config
property	否	String	语音合成特征字符串，组成形式为{language}_{speaker}_{domain}，即“语种_人员标识_领域”。发音人分为普通发音人和精品发音人，每次调用价格相同，针对精品发音人，每50字计一次调用，不足50字按一次计；普通发音人每100字计一次调用，不足100字按一次计。其中1个汉字、1个英文字母或1个标点均算作1个字符，具体费用请参考**价格计算器**。 ● 普通发音人：取值范围请参考**表 普通发音人property取值范围**。 ● 精品发音人：区域仅支持cn-north-4、cn-east-3，暂时不支持音高调节，取值范围请参考**表 精品发音人property取值范围**。使用精品发音人如果报错SIS.0411，请检查是否符合使用约束。 默认：chinese_xiaoyan_common，中文发音，小燕，温柔女声发音人。 父节点：config

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
speed	否	Integer	语速。 取值范围：-500~500 默认值：0 父节点：config **说明** 当取值为“0”时，表示一个成年人正常的语速，约为250字/分钟。设置该值时，语速和数值没有绝对的映射关系。
pitch	否	Integer	音高。 取值范围： -500~500 默认值：0 父节点：config
volume	否	Integer	音量。 取值范围：0~100 默认值：50 父节点：config

表 9-5 普通发音人 property 取值范围

名称	property取值	类型	使用场景	支持采样率（Hz）
小琪	chinese_xiaoqi_common	标准女声	客服	8k/16k
小雯	chinese_xiaowen_common	柔美女声	客服	8k/16k
小宇	chinese_xiaoyu_common	标准男声	电销	8k/16k
小夏	chinese_xiaoxia_common	热情女声	电销	8k/16k
小燕	chinese_xiaoyan_common	温柔女声	文学	8k/16k
小倩	chinese_xiaoqian_common	成熟女声	文学	8k/16k
小王	chinese_xiaowang_common	童声	童声	8k/16k
小呆	chinese_xiaodai_common	呆萌童声	童声	8k/16k
小婧	chinese_xiaojing_common	俏皮女声	新闻播报	8k/16k

版权所有 © 华为云计算技术有限公司

名称	property取值	类型	使用场景	支持采样率（Hz）
小宋	chinese_xiaosong_common	激昂男声	新闻播报	8k/16k
cameal	english_cameal_common	英文女声	英文	8k/16k

表 9-6 精品发音人 property 取值范围

名称	property取值	类型	使用场景	支持采样率（Hz）
华小夏	chinese_huaxiaoxia_common	热情女声	电销	8k/16k
华小唯	chinese_huaxiaowei_common	嗲柔女声	电销	8k/16k
华小颜	chinese_huaxiaoyan_common	严厉女声	电销	8k/16k
华晓阳	chinese_huaxiaoyang_common	朝气男声	电销	8k/16k
华晓刚	chinese_huaxiaogang_common	利落男声	客服	8k/16k
华小雯	chinese_huaxiaowen_common	柔美女声	客服	8k/16k
华小美	chinese_huaxiaomei_common	温柔女声	客服	8k/16k
华小飞	chinese_huaxiaofei_common	朝气男声	客服	8k/16k
华小荷	chinese_huaxiaohe_common	嘹亮女声	中英混合	8k/16k
华小叶	chinese_huaxiaoye_common	知性女声	中英混合	8k/16k
华小雪	chinese_huaxiaoxue_common	女童声	中英混合	8k/16k
华小娇	chinese_huaxiaojiao_common	成熟女声	纯中文	8k/16k
华小辉	chinese_huaxiaohui_common	男童声	中英混合	8k/16k
华小康	chinese_huaxiaokang_common	朝气男声	中英混合	8k/16k

版权所有 © 华为云计算技术有限公司

名称	property取值	类型	使用场景	支持采样率（Hz）
华小坤	chinese_huaxiaokun_common	成熟男声	中英混合	8k/16k
华小泉	chinese_huaxiaoquan_common	朝气男声	纯中文	8k/16k
华小璐	chinese_huaxiaolu_common	知性女声	新闻播报	8k/16k
华小靓	chinese_huaxiaoliang_common	嘹亮女声	新闻播报	8k/16k
华晓东	chinese_huaxiaodong_common	成熟男声	新闻播报	8k/16k
华小蕊	chinese_huaxiaorui_common	知性女声	中英混合	16k
华小萱	chinese_huaxiaoxuan_common	台湾女声	方言	8k/16k
华小闽	chinese_huaxiaomin_common	闽南女声	方言	8k/16k
华小舒	chinese_huaxiaoshu_common	舒缓女声	文学	8k/16k
华女侠	chinese_huanvxia_literature	武侠女生（只支持http形式调用）	文学	16k
华晓悬	chinese_huaxiaoxuan_literature	悬疑男声（只支持http形式调用）	文学	16k
华小龙	chinese_huaxiaolong_common	朝气男声	中英混合	16k
华小汝	chinese_huaxiaoru_common	柔美女声	中英混合	8k/16k
华小涵	chinese_huaxiaohan_common	知性女声	中英混合	8k/16k
华小宁	chinese_huaxiaoning_common	沉稳男声	中英混合	8k/16k
华小珍	chinese_huaxiaozhen_common	温柔女声	中英混合	8k/16k
华小曼	chinese_huaxiaoman_common	温柔女声	中英混合	16k

版权所有 © 华为云计算技术有限公司

名称	property取值	类型	使用场景	支持采样率（Hz）
华小芳	chinese_huaxiaofang_common	朝气女声	中英混合	16k
华小筠	chinese_huaxiaojun_common	成熟女声	中英混合	16k
alvin	english_alvin_common	成熟男声	纯英文	8k/16k
amy	english_amy_common	成熟女声	纯英文	8k/16k

响应参数

状态码： 200

表 9-7 响应 Body 参数

参数	是否必选	参数类型	描述
trace_id	否	String	服务内部的令牌，可用于在日志中追溯具体流程。 在某些错误情况下可能没有此令牌字符串。
result	否	**CustomResult** object	调用成功表示合成结果，调用失败时无此字段。

表 9-8 CustomResult

参数	是否必选	参数类型	描述
data	否	String	语音数据，以Base64编码格式返回。 用户如需生成音频，需要将Base64编码解码成byte数组，再保存为音频，音频格式同“audio_format”参数设置的值，默认为wav格式。 语音合成代码示例请参考**SDK**，SDK已对Base64转音频过程进行封装，可以直接获取音频文件。

状态码： 400

版权所有 © 华为云计算技术有限公司

表 9-9 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 调用语音合成API，将文本合称为语音，并调整语音的音色、语速、音高、音量

```
POST https://{endpoint}/v1/{project_id}/tts

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...

Request Body:
 {
  "text": "欢迎使用语音云服务。",
   "config":
   {
     "audio_format": "wav",
     "sample_rate": "8000",
     "property": "chinese_xiaoyan_common",
     "speed": 10,
     "pitch": 10,
     "volume": 60
   }
 }
```

- 使用Python3语言调用语音合成API，将文本合称为语音，并调整语音的音色、语速、音高、音量

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装requests，执行pip install requests
import requests
import json

def stts_demo():
    url = 'https://{{endpoint}}/v1/{{project_id}}/tts'  # endpoint和project_id需替换
    token = '用户对应region的token'
    text = '待识别的文本'
    header = {
        'Content-Type': 'application/json',
        'X-Auth-Token': token
    }
    body = {'text': text}
    resp = requests.post(url, data=json.dumps(body), headers=header)
    print(resp.text)

if __name__ == '__main__':
    stts_demo()
```

- 使用Java语言调用语音合成API，将文本合称为语音，并调整语音的音色、语速、音高、音量

```
import java.io.BufferedReader;
import java.io.InputStream;
import java.io.InputStreamReader;
import java.io.OutputStreamWriter;
```


版权所有 © 华为云计算技术有限公司

```
import java.net.HttpURLConnection;
import java.net.URL;

/**
 * 此demo仅供测试使用，强烈建议使用SDK
 */
public class SttsDemo {
  public void sttsDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      URL url = new URL("https://{{endpoint}}/v1/{{project_id}}/tts");
      String token = "对应region的token";
      HttpURLConnection connection = (HttpURLConnection)url.openConnection();
      connection.setRequestMethod("POST");
      connection.setDoInput(true);
      connection.setDoOutput(true);
      connection.addRequestProperty("Content-Type", "application/json");
      connection.addRequestProperty("X-Auth-Token", token);


      OutputStreamWriter osw = new OutputStreamWriter(connection.getOutputStream(), "UTF-8");
      String body = "{\"text\": \"123\"}";
      osw.append(body);
      osw.flush();
      InputStream is = connection.getInputStream();
      BufferedReader br = new BufferedReader(new InputStreamReader(is, "UTF-8"));
      while (br.ready()) {
        System.out.println(br.readLine());
      }

    } catch (Exception e) {
      e.printStackTrace();
    }

  }
  public static void main(String[] args) {
    SttsDemo sttsDemo = new SttsDemo();
    sttsDemo.sttsDemo();
  }
}
```

响应示例

状态码：200

成功响应示例

```
{
  "trace_id": "567e8537-a89c-13c3-a882-826321939651",
  "result":{
    "data": "/+MgxAAUeHpMAUkQAANhuRAC..."
  }
}
```

状态码：400

失败响应示例

```
{
    "error_code": "SIS.0032",
    "error_msg": "'audio_format' is invalid"
}
```

状态码

状态码请参见**状态码**。

版权所有 © 华为云计算技术有限公司

错误码

错误码请参见错误码。

9.2 SSML 标记语言介绍

本文为您介绍SSML（Speech Synthesis Markup Language）标记语言的功能、标签使用及示例。

概述

SSML 是一种基于 XML 的语音合成标记语言。使用 SSML 可以更加准确、具体的定义合成音频的效果，包括控制断句分词方式、发音、速度、停顿、声调和音量等特征。相比文本输入进行合成，有更好的操作灵活性。

说明

华为云语音合成服务的SSML实现，基于W3C的语音合成标记语言版本1.1，但并不支持W3C包含的所有的标记类型。目前华小飞，华小美，华小荷，华小叶，华小雪，华小娇，华小康，华小辉，华小坤，华小泉，华小龙，华小蕊，华小闽合成支持 SSML功能。该特性与语言无关。发音人property取值详见语音合成。

- 华小飞（朝气男声）：chinese_huaxiaofei_common，支持8k/16k采样率。
- 华小美（温柔女声）：chinese_huaxiaomei_common，支持8k/16k采样率。
- 华小荷（嘹亮男声）：chinese_huaxiaohe_common，支持8k/16k采样率。
- 华小叶（知性女声）：chinese_huaxiaoye_common，支持8k/16k采样率。
- 华小雪（女童声）：chinese_huaxiaoxue_common，支持8k/16k采样率。
- 华小娇（成熟女声，纯中文）：chinese_huaxiaojiao_common，支持8k/16k采样率。
- 华小辉（男童声）：chinese_huaxiaohui_common，支持8k/16k采样率。
- 华小康（朝气男声）：chinese_huaxiaokang_common，支持8k/16k采样率。
- 华小坤（成熟男声）：chinese_huaxiaokun_common，支持8k/16k采样率。
- 华小泉（朝气男声，纯中文）：chinese_huaxiaoquan_common，支持8k/16k采样率。
- 华小龙（朝气男声）：chinese_huaxiaolong_common，支持16k采样率。
- 华小蕊（知性女声）：chinese_huaxiaorui_common，支持16k采样率。
- 华小闽（闽南女生）：chinese_huaxiaomin_common，支持8k/16k采样率。

使用方式

语音合成服务支持的SSML标签请参见标签。详情请参考文档下方各标签的介绍和示例。语音合成的SSML功能支持多个 <say-as> 标签闭合嵌套于文本之中，如：

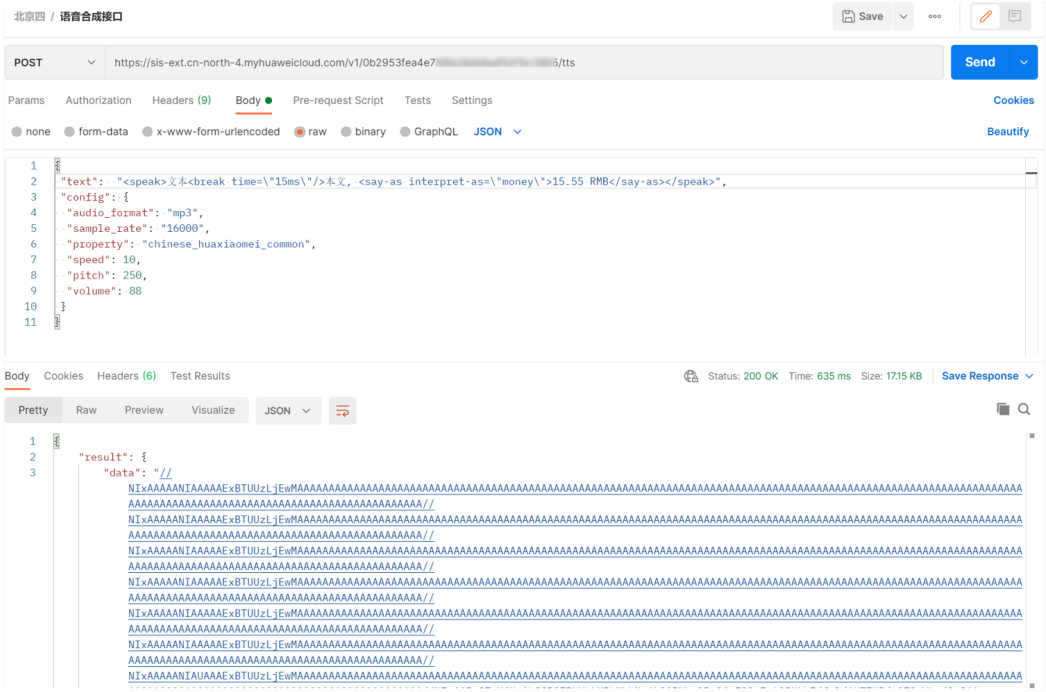
```
{
"text": "<speak>文本<break time=\"15ms\"/>本文, <say-as interpret-as=\"money\">15.55 RMB</say-as></speak>"
}
```

将带标签的文本作为text的参数值，上传至语音服务进行调用，以Java SDK为例：

版权所有 © 华为云计算技术有限公司

```
private String text = "<speak>文本<break time=\"15ms\"/>本文, <say-as interpret-as=\"money\">15.55 RMB</say-as></speak>";
1 usage
private String path = "D:\\ssml-test.mp3";
/**
 * 用于语音合成参数设置，例如发声人、音高、语速、音量、采样率、连接超时。所有参数均可以不设置，采用默认。
 *
 * @param request 语音合成请求
 */
1 usage
private void setParameter(TtsCustomRequest request) {
    // 设置语音格式，可选MP3，pcm等，默认wav
    request.setAudioFormat("mp3");
    // 音高，[-500, 500]，默认0
    request.setPitch(0);
    // 语速，[-500, 500]，默认0
    request.setSpeed(100);
    // 音量，[0, 100]，默认50
    request.setVolume(0);
    // 当前支持8000和16000，默认8000
    request.setSampleRate("16000");
    // 设置property，特征字符串，{language}_{speaker}_{domain}
```

使用postman调用格式如下：


北京四 / 语音合成接口
Save
POST
https://sis-ext.cn-north-4.myhuaweicloud.com/v1/0b2953fea4e7
5/tts
Send
Params
Authorization
Headers (9)
Body
Pre-request Script
Tests
Settings
Cookies
none
form-data
x-www-form-urlencoded
raw
binary
GraphQL
JSON
Beautify
"text": "<speak>文本<break time=\"15ms\"/>本文, <say-as interpret-as=\"money\">15.55 RMB</say-as></speak>",
"config": {
"audio_format": "mp3",
"sample_rate": "16000",
"property": "chinese_huaxiaomei_common",
"speed": 10,
"pitch": 250,
"volume": 88
}
Body
Cookies
Headers (6)
Test Results
Status: 200 OK
Time: 635 ms
Size: 17.15 KB
Save Response
Pretty
Raw
Preview
Visualize
JSON
"result": {
"data": "//
NIxAAAAANIAAAAAExBTUUzLjEwMAAAAA

2. <break>

描述：在任意位置，插入一个静音停顿

语法：

```
<!-- 空属性 -->
<break/>
<!-- 指定长度 -->
<break time="string"/>
<!-- 指定长度类 -->
<break strength="string"/>
```

属性：

表 9-10 标签属性

属性名称	属性类型	属性值	是否必选	描述
time	String	[number]s/[number]ms	否	以s/ms为单位设置静音持续时间(如："1s", "500ms")。 最大静音时间为10s。（华小龙，华小蕊不支持该属性）
strength	String	以下之一：none, x-weak, weak, medium,strong, x-strong	否	定义韵律（none=没有韵律，x-weak=很短，x-strong=很长）

标签关系：<break> 是空标签，不能包含任何标签。

示例：

```
一句话<break/>另外一句话
一句话<break time="20ms"/>另外一句话
一句话<break strength="strong"/>另外一句话
```

3. <phoneme>

描述:改变英文单词或中文多音字的发音。

语法：

```
<phoneme ph="string">这里输入字母</phoneme>
```

属性：

表 9-11 标签属性

属性名称	属性类型	属性值	是否必选	描述
ph	String	拼音，**CMU词典音素**	是	Pronunciation。

标签关系： <phoneme> 可以包含文本，不可以包含其他SSML标签。

示例：

版权所有 © 华为云计算技术有限公司

```
今天<phoneme ph="tian1 qi1">天气</phoneme>很好
The <phoneme ph="W EH1 DH AH0">weather</phoneme> is very good.
```

4. <say-as>

   描述:将文本解释为特定的类型内容，或者控制英文单词逐个字符拼写。

   语法：

```
<say-as interpret-as="string">这里输入字母</say-as>
```

   属性：

表 9-12 标签属性

属性名称	属性类型	属性值	是否必选	描述
iterpret-as	String	以下之一：money, date, figure,phone, english, spell; 具体含义参考下表。	是	将内容解释为给定类型。

表 9-13 iterpret-as 类型说明

类型	描述
money	钱
date	日期
figure	数次
phone	电话号码
english	英文单词
spell	逐个字母读英文

   标签关系： 可以包含文本 。

   示例：

```
今天<phoneme ph="tian1 qi1">天气</phoneme>很好
The <phoneme ph="W EH1 DH AH0">weather</phoneme> is very good.
```

```
<say-as interpret-as="money">15.55 RMB</say-as>
<say-as interpret-as="date">2022/3/8</say-as>
<say-as interpret-as="figure">175 厘米</say-as>
<say-as interpret-as="phone">151 12345678</say-as>
<say-as interpret-as="english">Hello</say-as>
<say-as interpret-as="spell">Hello</say-as>
```

5. <sub>

   描述：用其他词语代替合成文本中的词语。

   语法：

```
<sub alias="string">这里输入字母</sub>
```


版权所有 © 华为云计算技术有限公司

属性：

表 9-14 标签属性

属性名称	属性类型	属性值	是否必选	描述
alias	String	替代值	是	将标记的内容替换为此值。

标签关系： 可以包含文本 。

示例：

```
<sub alias="保罗">Paul</sub>是德国人
```


版权所有 © 华为云计算技术有限公司

10 实时语音合成接口

10.1 Websocket 握手请求

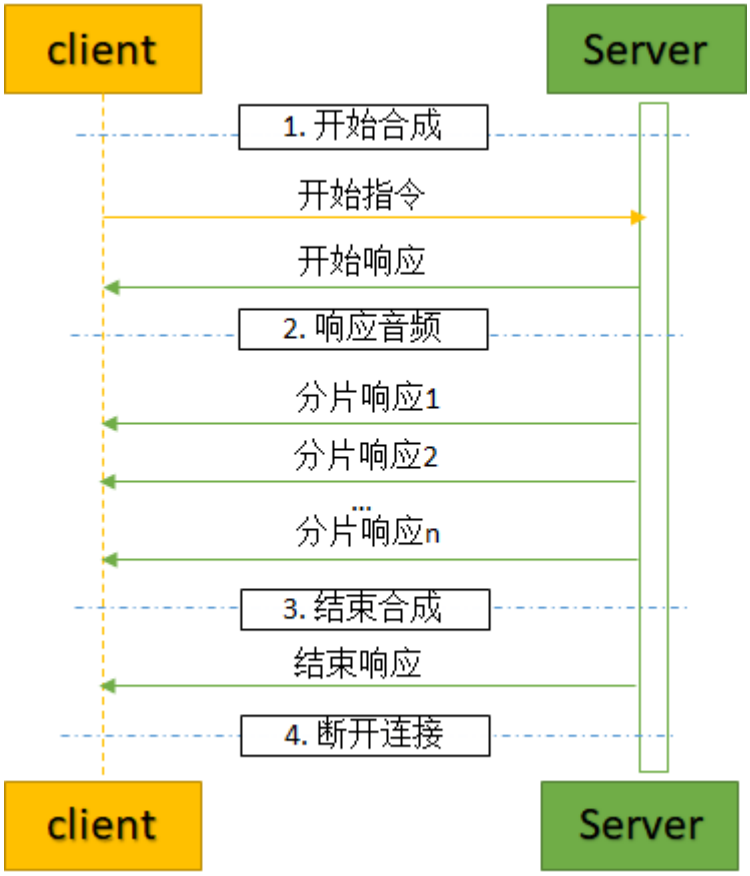
功能介绍

本接口提供实时流式语音合成。用户每次建立连接，发送待合成文本，服务端将合成结果响应给用户。一次连接只能发送一次文本，如果有多段文本需要合成，需要多次建立连接。实时语音合成和语音合成均属于语音合成接口，采用完全相同的计费规则，两者价格梯度可互相叠加，可参考价格计算器。

工作流程

从流程图可以看出，实时语音合成只需要发送1次开始请求，会陆续收到开始响应、合成音频的分片响应、结束响应。

版权所有 © 华为云计算技术有限公司



wss-URI

- wss-URI格式
  wss /v1/{project_id}/rtts
- 参数说明

表 10-1 参数说明

名称	是否必选	说明
project_id	是	项目编号。获取方法，请参见**获取项目ID**。

表 10-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
Enterprise-Project-Id	否	String	企业项目ID。SIS支持通过企业项目管理（EPS）对不同用户组和用户的资源使用，进行分账。 获取方法：进入“企业项目管理”页面，单击企业项目名称，在企业项目详情页获取Enterprise-Project-Id（企业项目ID）。 企业项目创建步骤请参见用户指南。 **说明** 账户创建企业项目后，在传参时，有以下三类场景。 ● 携带正确的ID，正常使用SIS服务，账单归到企业ID对应的企业项目中。 ● 携带错误的ID，正常使用SIS服务，账单的企业项目会被分类为“default”。 ● 不携带ID，正常使用SIS服务，账单的企业项目会被分类为“default”。

- 请求示例（伪码）

```
wss://{endpoint}/v1/{project_id}/rtts

Request Header:
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

Python3语言请求代码示例

```
# -*- coding: utf-8 -*-
# 此demo仅供测试使用，强烈建议使用sdk。需提前安装websocket-client, 执行pip install websocket-client
import websocket
import threading
import time
import json


def rtts_demo():
    url = 'wss://{{endpoint}}/v1/{{project_id}}/rtts'  # endpoint和project_id需替换
    text = '待合成文本'
    token = '用户对应region的token'
    header = {
        'X-Auth-Token': token
    }

    body = {
        'command': 'START',
        'text': text,
        'config': {
            'audio_format': 'pcm',
            'property': 'chinese_xiaoyu_common',
            'sample_rate': '8000'
        }
    }
```


版权所有 © 华为云计算技术有限公司

```
    def _on_message(ws, message):
        if isinstance(message, bytes):
            print('receive data length %d' % len(message))
        else:
            print(message)

    def _on_error(ws, error):
        print(error)

    ws = websocket.WebSocketApp(url, header, on_message=_on_message, on_error=_on_error)
    _thread = threading.Thread(target=ws.run_forever, args=(None, None, 30, 20))
    _thread.start()
    time.sleep(1)
    ws.send(json.dumps(body), opcode=websocket.ABNF.OPCODE_TEXT)
    time.sleep(10)
    ws.close()


if __name__ == '__main__':
    rtts_demo()
```

Java语言请求代码示例

```
import okhttp3.OkHttpClient;
import okhttp3.Request;
import okhttp3.Response;
import okhttp3.WebSocket;
import okhttp3.WebSocketListener;
import okio.ByteString;

/**
 * 此demo仅供测试使用, 强烈建议使用sdk
 * 使用前需已配置okhttp、okio jar包。jar包可通过下载SDK获取。
 */
public class RttsDemo {
  public void rttsDemo() {
    try {
      // endpoint和projectId需要替换成实际信息。
      String url = "wss://{{endpoint}}/v1/{{project_id}}/rtts";
      String token = "对应region的token";
      String text = "待合成文本";
      OkHttpClient okHttpClient = new OkHttpClient();
      Request request = new Request.Builder().url(url).header("X-Auth-Token", token).build();
      WebSocket webSocket = okHttpClient.newWebSocket(request, new MyListener());
      webSocket.send("{\"command\": \"START\", \"text\":\"" + text
          + "\", \"config\": {\"audio_format\": \"pcm\", \"property\": \"chinese_xiaoyu_common\"}}");
      Thread.sleep(10000);
      webSocket.close(1000, null);

    } catch (Exception e) {
      e.printStackTrace();
    }

  }

  class MyListener extends WebSocketListener {
    @Override

    public void onOpen(WebSocket webSocket, Response response) {
      System.out.println("conneected");
    }

    @Override
    public void onClosed(WebSocket webSocket, int code, String reason) {
      System.out.println("closed");
    }

    @Override
    public void onFailure(WebSocket webSocket, Throwable t, Response response) {
      t.printStackTrace();
```


版权所有 © 华为云计算技术有限公司

```
    }

    @Override
    public void onMessage(WebSocket webSocket, String text) {
      System.out.println(text);
    }

    public void onMessage(WebSocket webSocket, ByteString bytes) {
      byte[] data = bytes.toByteArray();
      System.out.println("receive data length is " + data.length);
    }

  }

  public static void main(String[] args) {
    RttsDemo rttsDemo = new RttsDemo();
    rttsDemo.rttsDemo();
  }
}
```

10.2 实时语音合成请求

10.2.1 开始语音合成请求

功能介绍

客户端与语音合成引擎建立Websocket连接后，可以发送开始语音合成请求进行语音合成。如果客户端基于该Websocket连接发送多次合成请求，需在每次请求重新建立websocket连接，一次连接只能处理一次合成请求。

请求消息

表 10-3 参数说明

名称	参数类型	是否必选	说明
command	String	是	需设置为START，表示开始识别请求。

版权所有 © 华为云计算技术有限公司

名称	参数类型	是否必选	说明
text	String	是	待合成的文本，chinese_huaxiaomei_common，chinese_huaxiaofei_common，chinese_huaxiaolong_common，chinese_huaxiaorui_common，chinese_huaxiaohe_common，chinese_huaxiaoye_common，chinese_huaxiaoxue_common，chinese_huaxiaojiao_common，chinese_huaxiaohui_common，chinese_huaxiaokang_common，chinese_huaxiaokun_common，chinese_huaxiaoquan_common，english_amy_common，english_alvin_common发音人支持长度不大于10000字符的文本，其他发音人支持长度不大于500字符的文本。
config	Object	否	配置信息。请参考**表10-4**。

表 10-4 config 数据结构

名称	参数类型	是否必选	说明
audio_format	String	否	语音格式头：pcm、alaw、ulaw、mp3。 默认：pcm
sample_rate	String	否	采样率：16000、8000赫兹。 默认：8000
property	String	否	可参考**表10-5**和**表10-6**。 默认：chinese_xiaoyan_common 实时语音合成和语音合成属于同一种资源，按次计费。实时语音合成普通发音人，每100字计一次。精品发音人每50字计一次。
speed	Integer	否	语速。 取值范围：-500~500 默认值：0
pitch	Integer	否	音高。 取值范围：-500~500 默认值：0

版权所有 © 华为云计算技术有限公司

名称	参数类型	是否必选	说明
volume	Integer	否	音量。 取值范围：0~100 默认值：50
subtitle	String	否	是否生成时间戳信息，不使用则不填写。 取值范围： “word_level”：文本级别时间戳。 “phoneme_level”：音素级别时间戳。 注：chinese_huaxiaomei_common、chinese_huaxiaofei_common、chinese_huaxiaohe_common、chinese_huaxiaoye_common、chinese_huaxiaoxue_common、chinese_huaxiaojiao_common、chinese_huaxiaohui_common、chinese_huaxiaokang_common、chinese_huaxiaokun_common、chinese_huaxiaoquan_common支持该功能，其他property暂不支持。

表 10-5 普通发音人 property 取值范围

名称	property取值	类型	使用场景	支持采样率（Hz）	支持audio_format
小琪	chinese_xiaoqi_common	标准女声	客服	8k/16k	pcm、alaw、ulaw
小宇	chinese_xiaoyu_common	标准男声	电销	8k/16k	pcm、alaw、ulaw
小燕	chinese_xiaoyan_common	温柔女声	文学	8k/16k	pcm、alaw、ulaw
小王	chinese_xiaowang_common	童声	童声	8k/16k	pcm、alaw、ulaw

版权所有 © 华为云计算技术有限公司

表 10-6 精品发音人 property 取值范围

名称	property取值	类型	使用场景	支持采样率（Hz）	支持audio_format
华小美	chinese_huaxiaomei_common	温柔女声	客服	8k/16k	pcm、mp3
华小飞	chinese_huaxiaofei_common	朝气男声	客服	8k/16k	pcm、mp3
华小汝	chinese_huaxiaoru_common	柔美女声	中英混合	8k/16k	pcm
华小涵	chinese_huaxiaohan_common	知性女声	中英混合	8k/16k	pcm
华小蕊	chinese_huaxiaorui_common	知性女声	中英混合	16k	pcm、mp3
华小龙	chinese_huaxiaolong_common	朝气男声	中英混合	16k	pcm、mp3
华小珍	chinese_huaxiaozhen_common	温柔女声	中英混合	8k/16k	pcm
华小荷	chinese_huaxiaohe_common	嘹亮女声	中英混合	8k/16k	pcm、mp3
华小叶	chinese_huaxiaoye_common	知性女声	中英混合	8k/16k	pcm、mp3
华小雪	chinese_huaxiaoxue_common	女童声	中英混合	8k/16k	pcm、mp3
华小娇	chinese_huaxiaojiao_common	成熟女声	纯中文	8k/16k	pcm、mp3
华小辉	chinese_huaxiaohui_common	男童声	中英混合	8k/16k	pcm、mp3
华小康	chinese_huaxiaokang_common	朝气男声	中英混合	8k/16k	pcm、mp3
华小坤	chinese_huaxiaokun_common	成熟男声	中英混合	8k/16k	pcm、mp3
华小泉	chinese_huaxiaoquan_common	朝气男声	纯中文	8k/16k	pcm、mp3
alvin	english_alvin_common	成熟男声	纯英文	8k/16k	pcm
amy	english_amy_common	成熟女声	纯英文	8k/16k	pcm

版权所有 © 华为云计算技术有限公司

示例

```
{
    "command": "START",
    "text": "欢迎使用语音云服务。",
    "config":
    {
        "audio_format": "pcm",
        "sample_rate": "8000",
        "property": "chinese_xiaoyan_common",
        "speed": 10,
        "pitch": 10,
        "volume": 60
    }
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

10.3 实时语音合成响应

10.3.1 开始合成响应

功能介绍

语音合成引擎收到实时语音合成请求时，首先向客户端发送合成开始响应，表示开始处理语音合成请求。

响应消息

表 10-7 响应参数

名称	参数类型	说明
resp_type	String	响应类型。参数值为START，表示开始语音合成。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。

示例

```
{
    "resp_type": "START",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651"
}
```


版权所有 © 华为云计算技术有限公司

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

10.3.2 语音合成结果响应

10.3.2.1 音频流数据

功能介绍

分多段返回二进制语音数据流，如果用户未设置语音格式，则默认返回pcm格式语音。

10.3.2.2 时间戳数据

功能介绍

实时语音合成服务在生成音频流的同时，可以生成每个汉字/英文单词的时间戳信息。该信息可用于视频字幕和驱动数字人口型。

参数设置

设置请求参数subtitle为"word_level"或"phoneme_level"时，开启时间戳功能。

响应信息

表 10-8 响应参数

名称	参数类型	说明
resp_type	String	响应类型。参数值为RESULT。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
result	List	时间戳信息

表 10-9 result 数据结构

名称	参数类型	说明
start_time	Integer	文本对应的合成音频的开始时间戳，单位是ms。
end_time	Integer	文本对应的合成音频的结束时间戳，单位是ms。

版权所有 © 华为云计算技术有限公司

名称	参数类型	说明
text	String	文本信息。
word_index	Integer	文本在整句中的位置，从0开始计数。
phonemes	List	音素时间戳信息，当subtitle取值为phoneme_level时，返回该信息。

表 10-10 phonemes 数据结构

名称	参数类型	说明
phoneme	String	音素文本信息。
start_time	Integer	音素对应的合成音频的开始时间戳，单位是ms。
end_time	Integer	音素对应的合成音频的结束时间戳，单位是ms。
phoneme_index	Integer	音素位置信息，从0开始。

示例

word_level

```
{
 'resp_type': 'RESULT',
 'trace_id':  'd34e3ccb-0383-4c76-a107-ec6ced44614f',
 'result':
        [
            {'start_time': 43980, 'end_time': 44210, 'word_index': 10, 'text': '从'},
            {'start_time': 44210, 'end_time': 45298, 'word_index': 11, 'text': '而'}
        ]
}
```

```
{
 'resp_type': 'RESULT',
 'trace_id':  'd34e3ccb-0383-4c76-a107-ec6ced44614f',
 'result':
        [
            {'start_time': 0, 'end_time': 384, 'text': 'Nice', 'word_index': 0},
            {'start_time': 384, 'end_time': 512, 'text': 'to', 'word_index': 1},
            {'start_time': 512, 'end_time': 800, 'text': 'meet', 'word_index': 2},
            {'start_time': 800, 'end_time': 1184, 'text': 'you.', 'word_index': 3},
            {'start_time': 1184, 'end_time': 1284, 'text': '', 'word_index': 4}
        ]
}
```

phoneme_level

```
{
    'resp_type': 'START',
    'trace_id': '22706fed-ab0e-4167-a6c7-e4dbc59898b1',
    'result':
            [
                {
                    'word_index': 217,
```


版权所有 © 华为云计算技术有限公司

```
                'text': '需',
                'start_time': 458
                'end_time': 650
                "phonemes": [
                    {'phoneme_index': 0, 'start_time': 458, 'end_time': 602, 'phoneme': 'x'},
                    {'phoneme_index': 1, 'start_time': 602, 'end_time': 650, 'phoneme': 'v1'}
                ]
            },
            {
                'word_index': 217,
                'text': '要',
                'start_time': 650
                'end_time': 922
                "phonemes": [
                    {'phoneme_index': 0, 'start_time': 650, 'end_time': 922, 'phoneme': 'iao4'}
                ]
            }
        ]
}
```

10.3.3 语音合成结束响应

功能介绍

当合成引擎处理完合成请求后，会发送合成结束响应。客户端收到该响应后关闭当前Websocket链接即可。

响应消息

表 10-11 响应参数

名称	参数类型	说明
resp_type	String	响应类型。参数值为END，表示语音合成结束。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
reason	String	结束原因。

表 10-12 结束原因表

名称	说明
NORMAL	正常结束。
ERROR	合成过程中发生错误。

示例

```
{
    "resp_type": "END",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "reason": "NORMAL"
}
```


版权所有 © 华为云计算技术有限公司

状态码

状态码请参见状态码。

错误码

错误码请参见错误码。

10.3.4 语音合成错误响应

功能介绍

合成引擎处理合成请求时如果发生错误则返回该响应。

响应消息

表 10-13 响应参数

名称	参数类型	说明
resp_type	String	响应类型。参数值为ERROR，表示错误响应。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
error_code	String	参见错误码列表。
error_msg	String	返回错误信息。

示例

```
{
    "resp_type": "ERROR",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "error_code":"SIS.0032",
    "error_msg":"'audio_format' is invalid"
}
```

状态码

状态码请参见状态码。

错误码

错误码请参见错误码。

10.3.5 严重错误响应

功能介绍

严重错误，通常指流程无法继续的情况。出现严重错误响应时，流程不再继续，服务器端会主动断连。

版权所有 © 华为云计算技术有限公司

响应消息

表 10-14 响应参数

名称	参数类型	说明
resp_type	String	响应类型。参数值为FATAL_ERROR，表示此次合成发生不可恢复的错误。
trace_id	String	服务内部的令牌，可用于在日志中追溯具体流程。
error_code	String	错误码列表。详细错误码解释，请参见错误码。
error_msg	String	返回错误信息。

示例

```
{
    "resp_type": "FATAL_ERROR",
    "trace_id": "567e8537-a89c-13c3-a882-826321939651",
    "error_code": "SIS.0304",
    "error_msg": "wait voice timeout"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

版权所有 © 华为云计算技术有限公司

11 热词管理接口

11.1 创建热词表

在一句话识别、实时语音识别服务中，如果在您的业务领域有一些特有的词，默认识别效果差的时候可以考虑使用热词管理功能，将这些词添加到词表，改善识别效果。

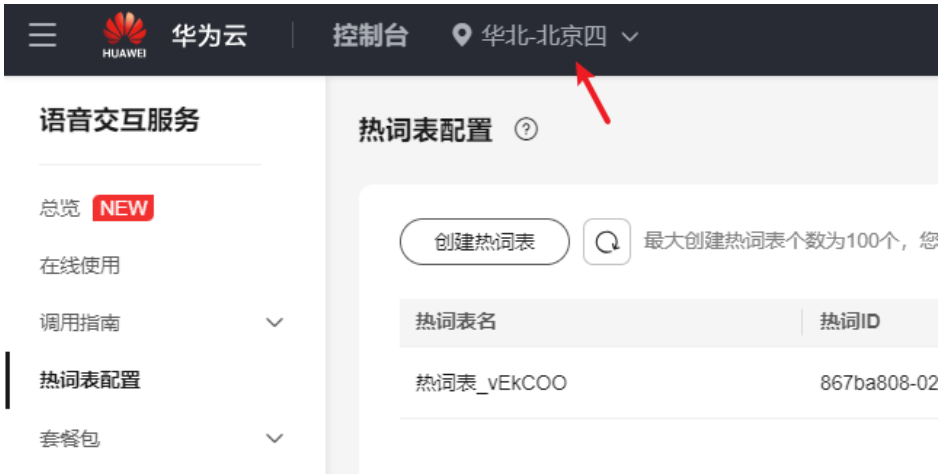
热词可以是姓名、公司名、某个领域的专有名词，如人名“王晓茗”（通常会识别为“王小明”）、运输域的“高栏”等。热词不建议太长，也不要有标点符号或者特殊符号。具体限制参见创建热词表更新热词表章节。

功能介绍

新建一个热词表，创建成功返回id。每个用户限制创建100个热词表。

调用热词时，需确保调用热词的区域与创建热词的区域一致。如下图位置可查询创建热词所在区域：

图 11-1 热词创建区域查询



调试

您可以在API Explorer中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

版权所有 © 华为云计算技术有限公司

约束与限制

- 所有含英文的热词均需大写，如：Eiffel铁塔（正确示例为：EIFFEL铁塔）。
- 所有数字均用相应的汉字表示，避免使用阿拉伯数字 ，如：12345服务热线（正确示例为：一二三四五服务热线）。
- 热词内容仅包含英文和中文，避免出现标点符号、特殊字符或空格，如：。，？.*等。
- 尽量避免配置单音节的英文单词，避免误召回，如：MAY/TEE等。
- 使用热词表功能时，同一个项目ID，可以共用热词表，不同的项目ID不能共用热词表。

URI

POST /v1/{project_id}/asr/vocabularies

表 11-1 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

请求参数

表 11-2 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

表 11-3 请求 Body 参数

参数	是否必选	参数类型	描述
name	是	String	热词表名，不可重复。内容限制为字母，数字，下中划线和井号，长度不超过32字符。
description	否	String	热词表描述，长度不超过255字符。
language	是	String	热词表语言类型。 language取值：chinese_mandarin，表示汉语普通话。

版权所有 © 华为云计算技术有限公司

参数	是否必选	参数类型	描述
contents	是	Array of String	支持中英混编热词，单个热词只能由英文字母和unicode编码的汉字组成，不能有其他符号，包括空格。 阿拉伯数字需写成汉字或英文（如“一”、“one”）。 单词库支持热词数上限10000。 中文单个热词长度上限32字符，英文为64字符。

响应参数

状态码： 200

表 11-4 响应 Body 参数

参数	是否必选	参数类型	描述
vocabulary_id	是	String	调用成功返回热词表ID，调用失败时无此字段。

状态码： 400

表 11-5 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

创建热词表

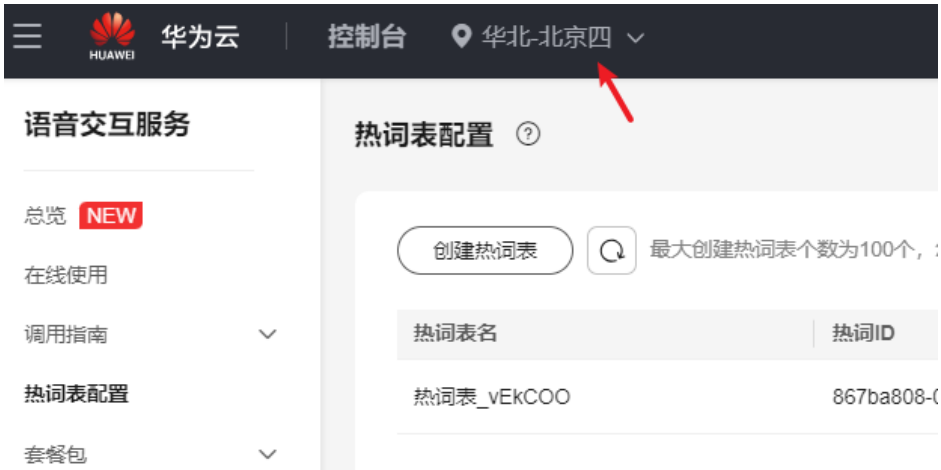
```
POST https://{endpoint}/v1/{project_id}/asr/vocabularies

Request Header:
Content-Type: application/json
X-Auth-Token: MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...

Request Body:
{
  "name": "telepower",
  "description": "telepower ",
```


版权所有 © 华为云计算技术有限公司

```
  "language": "chinese_mandarin",
  "contents": ["示例"]
}
```

调用热词时，需确保调用热词的区域与创建热词区域一致。如下图位置，可查看创建的热词表所在区域：

图 11-2 查看创建的热词表位置



响应示例

状态码： 200

成功响应示例

```
{
  "vocabulary_id": "CFD08A32-6176-4ad7-92F9-11ED015C8109",
}
```

状态码：400

失败响应示例

```
{
    "error_code":"SIS.0201",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

11.2 更新热词表

功能介绍

更新一个热词表，更新成功返回id。详细使用指导请参见**SIS服务使用简介**章节。

版权所有 © 华为云计算技术有限公司

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

PUT /v1/{project_id}/asr/vocabularies/{vocabulary_id}

表 11-6 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。
vocabulary_id	是	String	被更新的热词表id。

请求参数

表 11-7 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

表 11-8 请求 Body 参数

参数	是否必选	参数类型	描述
name	是	String	热词表名，不可重复。内容限制为字母，数字，下中划线和井号，长度不超过32字节。
description	否	String	热词表描述，长度不超过255字节。
language	是	String	热词表语言类型。 language取值：chinese_mandarin，表示汉语普通话。
contents	是	Array of strings	支持中英混编热词，单个热词只能由英文字母和unicode编码的汉字组成，不能有其他符号，包括空格。 单词库支持热词数上限10000。 单个热词长度上限32字节。

版权所有 © 华为云计算技术有限公司

响应参数

状态码： 200

表 11-9 响应 Body 参数

参数	是否必选	参数类型	描述
vocabulary_id	是	String	调用成功返回热词表ID，调用失败时无此字段。

状态码： 400

表 11-10 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 更新热词表

```
PUT https://{endpoint}/v1/{project_id}/asr/vocabularies/{vocabulary_id}

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...

Request Body:
{
  "name": "telepower",
  "description": "telepower 具体描述",
  "language": "chinese_mandarin",
  "contents": ["示例"]
}
```

响应示例

状态码： 200

成功响应示例

```
{
  "vocabulary_id": "CFD08A32-6176-4ad7-92F9-11ED015C8109",
}
```

状态码：400

失败响应示例

版权所有 © 华为云计算技术有限公司

```
{
    "error_code":"SIS.0201",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

11.3 查询热词表信息

功能介绍

通过热词表id查询热词表的信息和内容。详细使用指导请参见**SIS服务使用简介**章节。

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

GET /v1/{project_id}/asr/vocabularies/{vocabulary_id}

表 11-11 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。
vocabulary_id	是	String	被查询的热词表id。

请求参数

表 11-12 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

响应参数

状态码： 200

表 11-13 响应 Body 参数

参数	是否必选	参数类型	描述
vocabulary_id	是	String	调用成功返回热词表ID，调用失败时无此字段。
name	是	String	调用成功返回热词表名，调用失败时无此字段。
language	是	String	调用成功返回热词表语言类型，调用失败时无此字段。
description	是	String	调用成功返回热词表描述，调用失败时无此字段。
contents	是	Array of strings	调用成功返回热词列表，调用失败时无此字段。

状态码： 400

表 11-14 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 查询热词表信息

```
GET https://{endpoint}/v1/{project_id}/asr/vocabularies/{vocabulary_id}

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

响应示例

状态码： 200

成功响应示例

```
{
  "vocabulary_id": "CFD08A32-6176-4ad7-92F9-11ED015C8109",
```


版权所有 © 华为云计算技术有限公司

```
  "name": "telepower",
  "description": "telepower 具体描述",
  "language": "chinese_mandarin",
  "contents": ["示例"]
}
```

状态码：400

失败响应示例

```
{
    "error_code":"SIS.0201",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

11.4 删除热词表

功能介绍

通过热词表id删除热词表。详细使用指导请参见**SIS服务使用简介**章节。

查询

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

DELETE /v1/{project_id}/asr/vocabularies/{vocabulary_id}

表 11-15 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。
vocabulary_id	是	String	被删除的热词表id。

版权所有 © 华为云计算技术有限公司

请求参数

表 11-16 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

响应参数

无响应消息，以http状态码200表示成功。

状态码： 400

表 11-17 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

- 删除热词表

```
DELETE https://{endpoint}/v1/{project_id}/asr/vocabularies/{vocabulary_id}

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

响应示例

状态码：400

失败响应示例

```
{
  "error_code" : "SIS.0201",
  "error_msg" : "***"
}
```

状态码

状态码请参见**状态码**。

版权所有 © 华为云计算技术有限公司

错误码

错误码请参见**错误码**。

11.5 查询热词表列表

功能介绍

查询用户所有热词表列表。详细使用指导请参见**SIS服务使用简介**章节。

调试

您可以在**API Explorer**中调试该接口，支持自动认证鉴权。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能。

URI

GET /v1/{project_id}/asr/vocabularies

表 11-18 路径参数

参数	是否必选	参数类型	描述
project_id	是	String	项目编号。获取方法，请参见**获取项目ID**。

表 11-19 Query 参数

参数	是否必选	参数类型	描述
limit	否	Integer	每页查询显示的条目数量。 缺省值：**10**
offset	否	Integer	页码偏移量，表示从此页码偏移量开始查询，offset大于等于0。 缺省值：**0**

请求参数

表 11-20 请求 Header 参数

参数	是否必选	参数类型	描述
X-Auth-Token	是	String	用户Token。 用于获取操作API的权限。获取方法请参见**认证鉴权**。响应消息头中X-Subject-Token的值即为Token。

版权所有 © 华为云计算技术有限公司

表 11-21 请求 Body 参数

参数	是否必选	参数类型	描述
name	否	String	热词表的表名，用于筛选热词表表名。

响应参数

状态码： 200

表 11-22 响应 Body 参数

参数	是否必选	参数类型	描述
count	是	Integer	热词表数。
result	是	Array of **VocabInfo** objects	调用成功返回热词表列表，调用失败时无此字段。

表 11-23 VocabInfo

参数	是否必选	参数类型	描述
vocabulary_id	是	String	热词ID。
name	是	String	热词表名。
language	是	String	热词表语言类型。
description	是	String	热词表描述。

状态码： 400

表 11-24 响应 Body 参数

参数	参数类型	描述
error_code	String	调用失败时的错误码。 调用成功时无此字段。
error_msg	String	调用失败时的错误信息。 调用成功时无此字段。

请求示例

说明

“endpoint”即调用API的请求地址，不同服务不同区域的“endpoint”不同，具体请参见**终端节点**。

版权所有 © 华为云计算技术有限公司

- 获取热词表列表信息

```
GET https://{endpoint}/v1/{project_id}/asr/vocabularies

Request Header:
Content-Type: application/json
X-Auth-Token:
MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
```

响应示例

状态码：200

成功响应示例

```
{
  "result": [
    {
        "vocabulary_id": "5F85A74C-BED9-4a15-B66E-039251D877D6",
        "language": "chinese_mandarin",
        "name": "weather",
        "description": "no desc"
    },
    {
        "vocabulary_id": "50875954-7328-42ab-B236-B3EC6E22207A",
        "language": "chinese_mandarin",
        "name": "war",
        "description": "no desc"
    }
  ]
}
```

状态码：400

失败响应示例

```
{
    "error_code":"SIS.0201",
    "error_msg":"***"
}
```

状态码

状态码请参见**状态码**。

错误码

错误码请参见**错误码**。

版权所有 © 华为云计算技术有限公司

12 附录

12.1 示例音频

测试音频如表 示例音频所示，音频文件标题表示采样率和位宽。如8k16bit.pcm表示音频采样率为8k，位宽为16bit。

表 12-1 示例音频

音频格式	下载链接
mp3	https://sis-sample-audio.obs.cn-north-1.myhuaweicloud.com/16k16bit.mp3
wav	https://sis-sample-audio.obs.cn-north-1.myhuaweicloud.com/16k16bit.wav
pcm	https://sis-sample-audio.obs.cn-north-1.myhuaweicloud.com/16k16bit.pcm
pcm	https://sis-sample-audio.obs.cn-north-1.myhuaweicloud.com/8k16bit.pcm

12.2 获取项目 ID

从控制台获取项目 ID

1. 登录管理控制台。
2. 在控制台中，鼠标移动至右上角的用户名处，在下拉列表中选择“我的凭证”。

版权所有 © 华为云计算技术有限公司

图 12-1 我的凭证入口

3. 在“我的凭证”页面，可以查看用户名、账号名，在项目列表中查看项目。

图 12-2 查看项目 ID

多项目时，展开“所属区域”，从“项目ID”列获取子项目ID。

调用 API 获取项目 ID

获取项目ID的接口为“GET https://{Endpoint}/v3/projects”，其中{Endpoint}为IAM的终端节点。接口的认证鉴权请参见**认证鉴权**。

响应示例如下，例如，语音交互服务部署的区域为“cn-north-4”，响应消息体中查找“name”为“cn-north-4”，其中projects下的“id”即为项目ID。

```
{
    "projects": [
        {
            "domain_id": "65382450e8f64ac0870cd180d14e684b",
            "is_domain": false,
            "parent_id": "65382450e8f64ac0870cd180d14e684b",
            "name": "project_name",
            "description": "",
            "links": {
                "next": null,
                "previous": null,
                "self": "https://support-intl.huaweicloud.com/zh-cn/devg-apisign/api-sign-
provide.htmlcd05f897d6b99"
            },
            "id": "a4a5d4098fb4474fa22cd05f897d6b99",
            "enabled": true
        }
    ],
    "links": {
        "next": null,
        "previous": null,
        "self": "https://www.example.com/v3/projects"
    }
}
```

12.3 获取账号 ID

在调用接口的时候，部分URL中需要填入账号ID（domain-id），所以需要先在管理控制台上获取到账号ID。账号ID获取步骤如下：

1. 注册并登录管理控制台。
2. 在控制台中，鼠标移动至右上角的用户名处，在下拉列表中选择“我的凭证”。

图 12-3 我的凭证入口

3. 在“我的凭证”页面的项目列表中查看账号ID。

图 12-4 查看账号 ID

12.4 获取 AK/SK

如果已生成过AK/SK，则可跳过此步骤，找到原来已下载的AK/SK文件，文件名一般为：credentials.csv。

如下图所示，文件包含了租户名（User Name），AK（Access Key Id），SK（Secret Access Key）。

图 12-5 credential.csv 文件内容

	A	B	C
1	User Name	Access Key Id	Secret Access Key
2	hu[redacted]dg	QTWA[redacted]UT2QVKYUC	MFyfvK41ba2[redacted]npdUKGpownRZlmVmHc

AK/SK生成步骤：

1. 登录管理控制台。
2. 单击右上角的用户名，在下拉列表中单击“我的凭证”。
3. 单击“访问密钥”。
4. 单击“新增访问密钥”。

版权所有 © 华为云计算技术有限公司

5. 依据界面提示，获取并下载秘钥，请妥善保管。

图 12-6 获取访问秘钥

12.5 公共请求参数

公共请求参数如**表 公共请求参数**所示。

表 12-2 公共请求参数

参数	是否必选	说明	示例
Content-type	是	发送的实体的MIME类型。	application/json
Content-Length	POST/PUT请求必填。GET不能包含。	请求body长度，单位为Byte。	3495
Upgrade	发送WebSocket握手请求时，包含该头域。	upgrade是HTTP1.1中用于定义转换协议的header域。	Websocket
Connection	发送WebSocket握手请求时，包含该头域。	带有Upgrade头的HTTP1.1消息须含有Connection。	Upgrade
Sec-WebSocket-Key	发送WebSocket握手请求时，包含该头域。	采用Base64编码的随机16字节长的字符序列。用来发送给服务器使用，服务器会使用此字段组装成另一个key值放在握手返回信息里发送客户端。	x3JJHMbDL1EzLkh9GBhXDw==
Sec-WebSocket-Version	发送WebSocket握手请求时，包含该头域。	标识了客户端支持WebSocket协议的版本。	13

参数	是否必选	说明	示例
X-Auth-Token	是	用户Token。	MIINRwYJKoZIhvcNAQcCoIINODCCDTQCAQExDTALBglghkgBZQMEAgEwgguVBgkqhkiG...
X-Language	请求语言类型。	否，默认为zh-cn。	en-us

说明

- 其它header属性，请遵照https协议。
- 调用业务接口时，在请求消息头中需增加Content-Type的消息体。

12.6 公共响应参数

公共响应参数如表 公共响应参数所示。

表 12-3 公共响应参数

参数名	说明
Content-Length	响应消息体的字节长度，单位为Byte。
Date	系统响应的时间。
Content-type	发送的实体的MIME类型，取值为application/json。
Upgrade	发送WebSocket握手请求时，响应消息包含该头域，内容为Websocket。
Connection	发送WebSocket握手请求时，响应消息包含该头域，内容为Upgrade。
Sec-WebSocket-Accept	结合Sec-WebSocket-Key提供基本的防护，比如恶意的连接，或者无意的连接。

12.7 状态码

表 12-4 状态码

状态码	说明
100	Continue
101	Switching Protocols

状态码	说明
200	OK
201	Created
202	Accepted
203	Non-Authoritative Information
204	NO Content
205	Reset Content
206	Partial Content
300	Multiple Choices
301	Moved Permanently
302	Found
303	See Other
304	Not Modified
305	Use Proxy
306	Unused
400	Bad Request
401	Unauthorized
402	Payment Required
403	Forbidden
404	Not Found
405	Method Not Allowed
406	Not Acceptable
407	Proxy Authentication Required
408	Request Timeout
409	Conflict
410	Gone
411	Length Required
412	Precondition Failed
413	Request Entity Too Large
414	Request URI Too Long
415	Unsupported Media Type

状态码	说明
416	Requested Range Not Satisfiable
417	Expectation Failed
422	Unprocessable Entity
429	Too Many Requests
500	Internal Server Error
501	Not Implemented
502	Bad Gateway
503	Service Unavailable
504	Gateway Timeout
505	HTTP Version Not Supported

12.8 错误码

调用接口出错后，将不会返回结果数据。调用方可根据每个接口对应的错误码来定位错误原因。当调用出错时，HTTPS请求返回一个 4xx 或 5xx 的 HTTPS状态码。返回的消息体中是具体的错误代码及错误信息。在调用方找不到错误原因时，可以联系华为云客服，并提供错误码，以便尽快帮您解决问题。

错误响应 Body 体格式说明

当接口调用出错时，会返回错误码及错误信息说明，错误响应的Body体格式如下所示。

```
{

    "error_code": "SIS.0032",
    "error_msg": "'audio_format' is invalid"
}
```

其中，error_code表示错误码，error_msg表示错误描述信息。

错误码说明

当您调用API时，如果遇到“APIGW”开头的错误码，请参见**API网关错误码**进行处理。

返回错误码	描述	处理措施
SIS.0001	连接错误，如连接obs错误或者连接redis错误。	请联系技术支持工程师。
SIS.0003	token缺少用户信息或者委托已创建或者内部错误。	请联系技术支持工程师。

返回错误码	描述	处理措施
SIS.0100	服务内部错误。	请联系技术支持工程师。
SIS.0101	验证Token异常。	● 请检查请求消息头中的X-Auth-Token是否正确的填写，是否缺少字符。 ● token请求参数是否正确填写。 ● projectId在url是否准确填写。 ● 账号是否被冻结。 ● 账户是否欠费。 ● 获取token时，请求中需添加scope参数，获取项目级别token，调用sis时请勿使用全局token。
SIS.0102	鉴权信息缺失。	请检查请求消息头中的X-Auth-Token是否存在。
SIS.0103	实名认证缺失。	请进行账号实名认证。
SIS.0012	请求体中字段缺失。	请检查请求参数是否缺失必须字段。
SIS.0031	请求参数不支持。	请检查请求参数是否正确。

版权所有 © 华为云计算技术有限公司

返回错误码	描述	处理措施
SIS.0032	请求体中的json不正确。	请检查请求体中的json是否正确。 ● 确保设置的audio_format格式的值是合法，待识别的音频格式和audio_format参数格式要保证一致。对于录音文件识别接口，如果不确定该参数取值，可改为auto进行调试。 ● 使用实时语音识别时请确认是否分片发送音频，并在wss握手后发送start命令。 ● 报错data url is empty，时请参考**配置OBS访问权限**，设置OBS权限，并检查使用的URL是否正确。 ● 报错data base64 encode invalid，请检查语音转Base64后的字符串是否存在错误，如多空格、字符错误等。 ● 报错xx cannot be empty，表示该字段不可为空，需要赋值。例如创建热词请求时报错language cannot be empty，则表示请求body体中language字段不可为空。 ● 报错callback_url is invalid or too long，请检查使用的URL格式是否正确。
SIS.0022	产品不可购买。	产品暂时不开放购买，请联系技术支持工程师。
SIS.0023	产品购买失败。	产品购买失败，请联系技术支持工程师。
SIS.0024	不允许更新限制。	产品暂时不允许更新限制，请联系技术支持工程师。
SIS.0033	引擎响应超时。	请联系技术支持工程师。
SIS.0201	热词表未找到。	请检查请求参数是否异常，或联系技术支持工程师。
SIS.0203	部分热词太长或非法。	请根据错误信息，修改不合法的热词。
SIS.0204	热词表参数名已存在。	请修改热词表名称。
SIS.0205	语言不支持。	请修改热词表语言。
SIS.0206	保存热词失败。	请联系技术支持工程师。

返回错误码	描述	处理措施
SIS.0207	热词内容损坏。	请联系技术支持工程师。
SIS.0208	热词表数量过多。	请删除不需要的热词表，如有配置更多热词表的需求请联系客服人员。
SIS.0301	输入的audio_format参数不匹配模型。	请检查请求参数是否正确。
SIS.0302	内部服务异常。	请联系技术支持工程师。
SIS.0303	引擎连接失败。	请重试，多次出现请联系专业工程师。
SIS.0304	音频等待超时。	客户端长时间没有发送语音，服务端超过20s没有收到语音会产生该报错。 ● 请减小数据发送间隔。 ● 检查代码是否存在发送音频后睡眠过多时间。 ● 检查是否发送end请求，是否关闭客户端。
SIS.0305	语音识别异常。	请重试，或联系技术支持工程师。
SIS.0306	语音识别异常。	请重试，或联系技术支持工程师。
SIS.0307	语音识别异常。	请重试，或联系技术支持工程师。
SIS.0309	实时语音识别音频时长超过最大值。	请注意音频的时长限制，流式一句话（short-stream）和单句模式（sentence-stream）最长识别音频长度为60秒，连续模式（continue-stream）最长识别音频长度为5小时。
SIS.0312	超过最大资源并发限制。	如果应用程序发送请求频率过高，请降低请求频率，或联系技术支持工程师。
SIS.0401	语音合成异常。	请重试，或联系技术支持工程师。
SIS.0402	语音合成输入的sample_rate参数非法。	请检查语音合成sample_rate采样率与property音色取值范围支持的采样率是否匹配。
SIS.0410	语音合成输入的audio_format参数非法。	请检查请求参数是否正确。
SIS.0411	语音合成输入的property参数非法。	请检查请求参数是否正确。
SIS.0412	连接语音合成引擎超时。	请重试，或联系技术支持工程师。

返回错误码	描述	处理措施
SIS.0413	语音合成内部错误。	请重试，或联系技术支持工程师。
SIS.0414	语音合成等待语音合成超时。	请重试，或联系技术支持工程师。
SIS.0415	语音合成请求体参数错误。	请检查请求体是否准确。
SIS.0416	等待用户发送合成文本超时。	请重新发送需要合成的文本内容。
SIS.0417	TTS引擎错误。	请重试，或联系技术支持工程师
SIS.0418	当前链接已有合成任务。	请勿重复发送开始合成命令。
SIS.0419	非法ssml。	输入文本内容包含非法字符。
SIS.0504	录音文件识别极速版待识别音频无效。	请检查音频文件大小是否满足要求。
SIS.0506	找不到代理错误。	● 请检查网络是否有设置代理。 ● 请在控制台检查OBS是否已授权。
SIS.0507	录音文件识别极速版输入参数无效。	请检查obs_bucket_name和obs_object_key参数是否为空。
SIS.0511	录音文件识别极速版待识别音频无效。	请检查音频时长是否满足要求。
SIS.0512	录音文件识别极速版输入参数无效。	请检查obs_bucket_name和obs_object_key对应的桶名和文件名是否存在。
SIS.0513	录音文件识别极速版输入参数无效。	请检查obs_object_key参数是否非法。
SIS.0533	录音文件识别请求参数错误。	请检查请求参数是否正确。
SIS.0534	录音文件识请求体错误。	请检查请求体是否正确。 ● data url is unreachable报错，请**配置OBS访问权限**。 ● data syntax error！报错请检查数据格式或编码格式是否符合要求。 ● asr data url param is invalid报错，请使用存放在OBS的音频文件路径，并确保OBS的region要和请求服务的region保持一致。

版权所有 © 华为云计算技术有限公司

返回错误码	描述	处理措施
SIS.0535	录音文件识请求的文件类型不支持。	• 确保设置的audio_format格式的值是合法。录音文件识别支持的音频格式如表7-7所示。 • 确保待识别音频的格式和接口参数中设置的audio_format参数格式一致。
SIS.0536	录音文件识任务提交数量达到限制。	请稍后重试。
SIS.0537	录音文件太大。	请尝试减少文件大小，比如分割为多个。
SIS.0538	录音文件识任务超时。	请重试。
SIS.0540	录音文件大小。	请尝试增大录音文件大小。
SIS.0541	录音文件识别 property 和 format 参数不匹配。	请检查Property 和 format 格式是否正常。
SIS.0601	短语音识别或语音合成输入参数不合法。	请检查请求参数是否正确和完整。
SIS.0602	短语音识别语音格式不支持。	请检查提交的语音编码格式是否在支持的范围内。
SIS.0604	短语音识别文件大小不符合要求或语音合成字数超过上限。	请尝试减小语音文件的大小或减少请求合成字数，注意使用OBS链接提交时，经过Base64编码之后文件大小会增加。
SIS.0605	短语音识别或语音合成出现内部错误。	请重试，或联系技术支持工程师。
SIS.0608	URL非法。	请检查URL，注意必须使用对应Region的OBS链接。
SIS.0609	从指定URL下载音频文件失败。	请检查OBS是否为公共读状态，或开通了委托授权。
SIS.0701	输入的音频/视频/试题文本大小超过限制。	请参考错误信息和API资料，减少音频/视频/试题文本大小。
SIS.0702	输入参数不合法。 例如，评测语言/评测模式不支持；音频/视频格式不支持；试题文本无效等情况。	请参考错误信息和API资料，更换正确的输入参数。
SIS.30003	下载OBS上的音频失败。	请检查OBS上音频是否可访问。权限配置方法请参见配置OBS访问权限。
SIS.30004	下载热词文件失败。	请检查热词是否存在。

版权所有 © 华为云计算技术有限公司

返回错误码	描述	处理措施
SIS.30006	音频解码失败。	请检查音频格式是否正常，音频是否是空文件。
SIS.30007	转写引擎加载热词失败。	请检查热词是否存在。
SIS.30008	转写引擎转写失败。	请检查参数或联系技术支持工程师。

12.9 配置 OBS 访问权限

OBS 服务

OBS全称Object Storage Service（对象存储服务），提供海量、安全、高可靠、低成本的数据存储能力，可供用户存储任意类型和大小的数据。

EI企业智能服务对于图片、语音等多媒体文件支持直接使用华为云OBS服务的数据处理方式，以减少服务使用成本，降低服务的响应时长，提升服务使用的体验。语音交互服务优先推荐使用同一region下OBS的音频链接作为传入音频的URL。

考虑到数据的安全，语音交互服务无法直接获取到用户OBS数据，需要用户开启公共读授权或者语音交互服务授权。

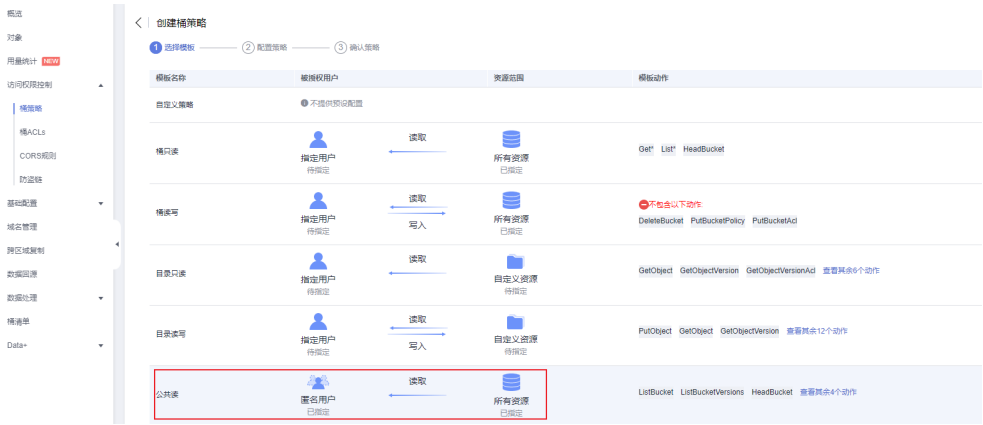
开启公共读授权

开启公共授权，则数据全网可见，所有用户均可访问，不推荐使用此方法，该方法仅针对同一region服务有效。

配置公共读可参考**配置标准桶策略**，创建桶策略为“**公共读**”。一般私密数据不建议用此方法。

目前仅支持访问用户个人OBS下的音频的链接，不支持读取其他用户公共读的链接。

图 12-7 创建“公共读”桶策略



开启语音交互服务授权

开启公共读授权访问，对于敏感信息，如个人私有数据，存在泄漏风险。可考虑开启语音交互服务授权。

进入语音交互服务，单击“服务授权”，选择确认对象存储服务授权。

主账号或者有Security Administrator权限的子账号才可以创建委托。

上传音频

1. 在OBS“桶列表”页，单击创建的OBS桶。

图 12-8 进入 OBS 桶

2. 进入“对象”页面，将音频数据上传至OBS桶中。

图 12-9 上传视频数据

获取音频 URL

1. 单击已上传的数据名称，进入数据详情页。

版权所有 © 华为云计算技术有限公司

图 12-10 进入数据详情页

2. 复制链接，获取数据URL。

图 12-11 获取数据 URL

版权所有 © 华为云计算技术有限公司