

解决方案实践

# 语音识别解决方案

文档版本 1.0  
发布日期 2023-12-05



版权所有 © 华为技术有限公司 2024。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。

# 安全声明

## 漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

---

# 目录

---

<b>1 方案概述</b> .....	<b>1</b>
<b>2 资源和成本规划</b> .....	<b>3</b>
<b>3 实施步骤</b> .....	<b>4</b>
3.1 准备工作.....	4
3.2 快速部署.....	12
3.3 开始使用.....	17
3.4 快速卸载.....	18
<b>4 附录</b> .....	<b>20</b>
<b>5 修订记录</b> .....	<b>21</b>

# 1 方案概述

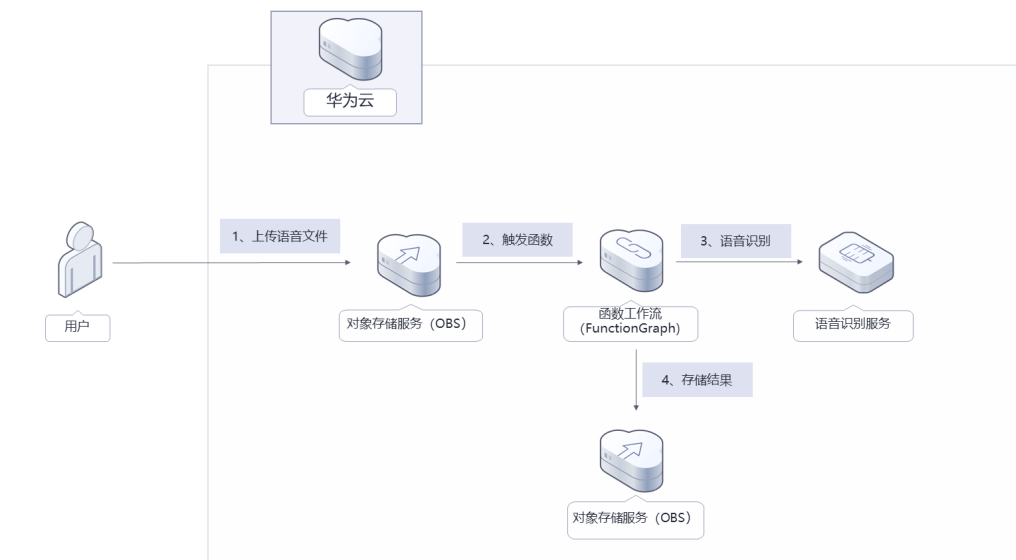
## 应用场景

该解决方案基于华为云语音交互服务语音识别构建，可自动将用户上传到对象存储服务的wav语音文件转化为文字，并将结果存放到指定OBS桶。该方案可以将用户上传在OBS的语音文件识别成可编辑的文本，支持中文普通话的识别和合成，其中语音识别还支持带方言口音的普通话识别以及方言（四川话、粤语和上海话）的识别。适用于如下场景：识别客服、客户的语音，进一步通过文本检索，检查有没有违规、敏感词、电话号码等信息。对会议记录的音频文件，进行快速的识别，转化成文字，方便进行会议记录等场景。

## 方案架构

该解决方案部署架构如下图所示：

图 1-1 方案架构



部署该方案中需要使用的资源：

- 创建2个对象存储服务OBS桶，一个用于用户上传和存储用户的语音文件；另一个用于存储语音识别结果，用户可以从该桶中获取结果并处理。

- 函数工作流FunctionGraph，只需编写业务函数代码并设置运行的条件，即可以弹性、免运维、高可靠的方式运行。
- 语音识别服务，将wav语音文件转化为文字。

## 方案优势

- 高识别  
该方案基于深度学习技术，对特定领域场景的语音识别进行优化，识别率高。
- 稳定可靠  
该方案成功应用于各类场景，基于华为等企业客户的长期实践，经受过复杂场景考验。
- 可定制化  
该方案针对客户的特定场景需求，定制垂直领域的语音识别模型，识别效果更精确。

## 约束与限制

- 部署该解决方案前，您需注册华为账号并开通华为云，完成实名认证，且账号不能处于欠费或冻结状态，如使用包周期部署确保余额充足。
- 部署该解决方案前，您需要按照[3.1 准备工作](#)完成相应服务开通及授权，上传的语音文件时长不超过30分钟，文件大小不超过100M。

# 2 资源和成本规划

该解决方案主要部署如下资源，不同产品的花费仅供参考，实际以收费账单为准，具体请参考华为云[官网价格](#)：

表 2-1 资源和成本规划

华为云服务	配置示例	每月花费
录音文件识别服务	<ul style="list-style-type: none"><li>区域：华北-北京四</li><li>计费模式：包年包月套餐包</li><li>计费价格：15元/10h</li></ul>	15元
对象存储服务	<ul style="list-style-type: none"><li>区域：华北-北京四</li><li>数据存储（多AZ存储）：0.1390元/GB/月</li><li>请求费用：0.0100元/万次</li></ul>	预计每月新增1GB数据量，花费0.1390元。 该方案存储费用消耗较低，详细请参考每月账单。
函数 workflow	<ul style="list-style-type: none"><li>区域：华北-北京四</li><li>请求次数0-100万次/月：免费</li><li>计量时间：0-400,000 GB/秒：免费</li><li>节点执行次数：0-5000次：免费</li></ul>	0
合计		15.14元

# 3 实施步骤

- 3.1 准备工作
- 3.2 快速部署
- 3.3 开始使用
- 3.4 快速卸载

## 3.1 准备工作

该解决方案部署，需要开通、配置以下授权。

### 购买、配置语音交互服务录音文件识别

步骤1 登录[语音交互服务](#)，选择录音文件识别，单击购买录音识别套餐包。

图 3-1 登录语音交互服务

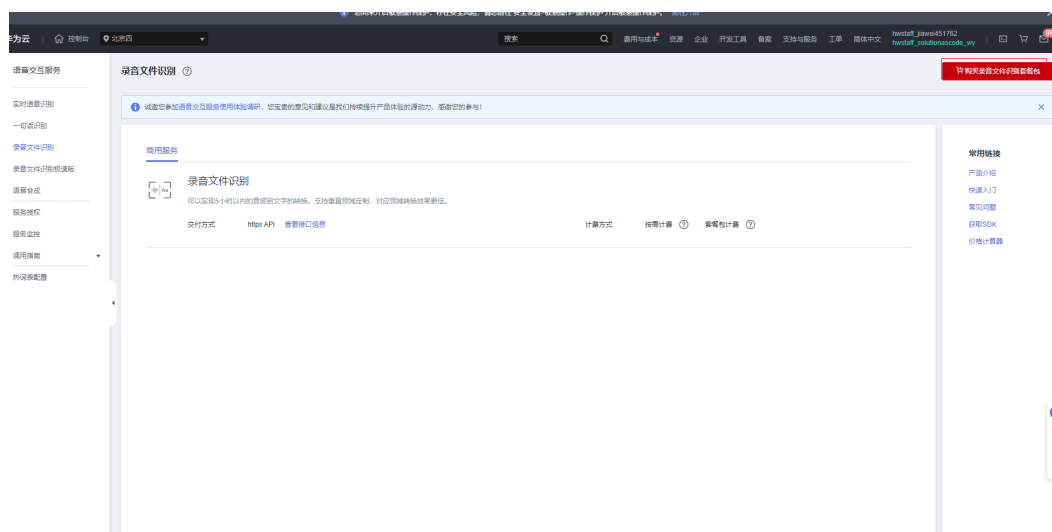
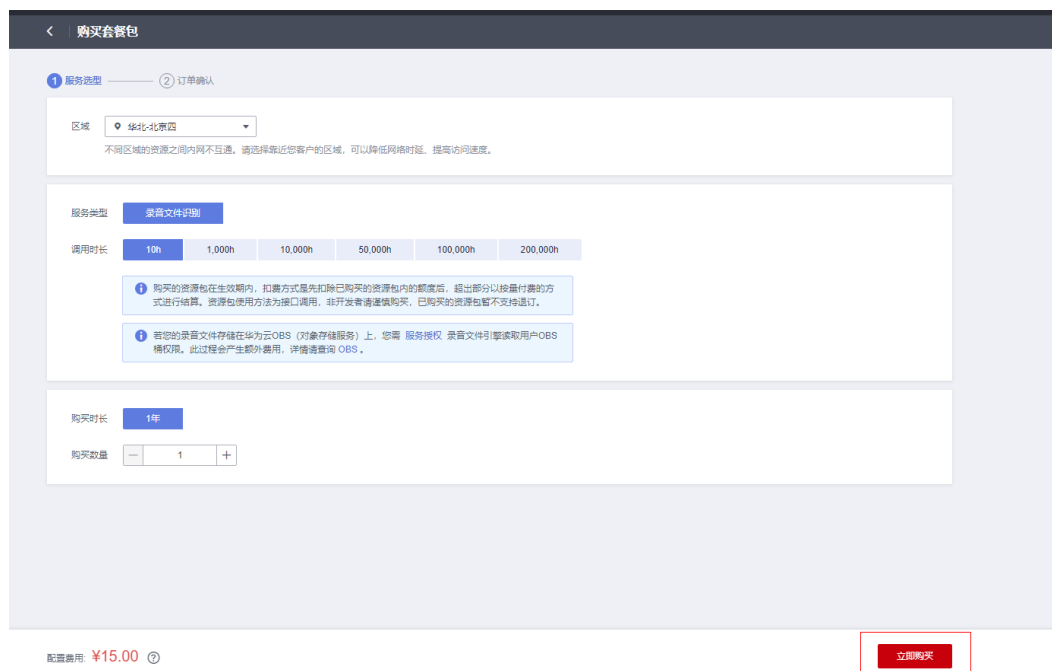


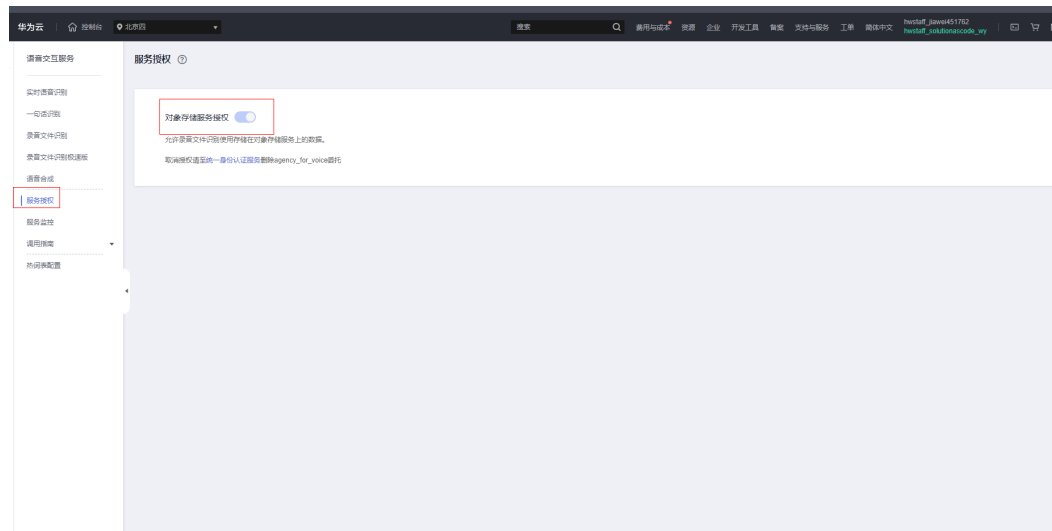


图 3-2 购买录音套餐



步骤2 选择服务授权，开通对象服务授权。

图 3-3 开通对象服务授权



----结束

## 创建 rf\_amdin\_trust 委托

步骤1 登录[华为云控制台](#)，打开“统一身份认证”菜单。

图 3-4 统一身份认证菜单



步骤2 进入“委托”菜单，搜索“rf\_admin\_trust”委托。

图 3-5 委托列表



- 如果委托存在，则不用执行接下来的创建委托的步骤
- 如果委托不存在时执行接下来的步骤创建委托

步骤3 单击步骤2界面中的“创建委托”按钮，在委托名称中输入“rf\_admin\_trust”，选择“普通账号”，委托的账号，输入“op\_svc\_IAC”，单击“下一步”。

图 3-6 创建委托



步骤4 在搜索框中输入” Tenant Administrator” 权限，并勾选搜索结果。

图 3-7 选择策略



步骤5 选择“所有资源”，并单击下一步完成配置。

图 3-8 设置授权范围



步骤6 “委托”列表中出现“rf\_admin\_trust”委托则创建成功。

图 3-9 委托列表



----结束

## 创建 IAM Agency Management FullAcces 策略

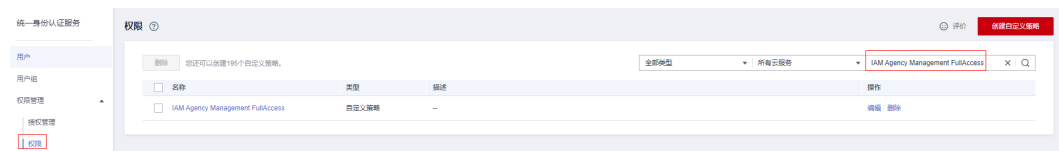
步骤1 打开“统一身份认证”菜单。

图 3-10 统一身份认证菜单



步骤2 进入“权限管理”->“权限”菜单，在搜索框输入“IAM Agency Management FullAcces”当前账号是否存在IAM委托管理权限。

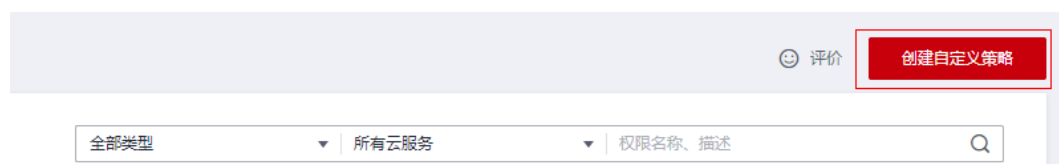
图 3-11 权限列表



- 如果搜索结果不为空，则当前账号已经存在IAM委托管理权限，不需要重复创建
- 如果过搜索结果为空，则继续创建“IAM Agency Management FullAcces”权限

步骤3 单击“创建自定义策略”按钮。

图 3-12 创建自定义策略



步骤4 输入策略名称为“IAM Agency Management FullAcces”，选择“JSON视图”，在策略内容中输入如下JSON代码，单击确认按钮。

图 3-13 创建自定义策略

策略名称: IAM Agency Management FullAccess

策略配置方式: 可视化视图 | JSON视图

策略内容:

```
1 {
2   "Version": "1.1",
3   "Statement": [
4     {
5       "Action": [
6         "iam:agencies:createAgency",
7         "iam:agencies:listAgencies",
8         "iam:agencies:getAgency",
9         "iam:agencies:deleteAgency",
10        "iam:agencies:updateAgency",
11        "iam:permissions:revokeRoleFromAgencyOnProject",
12        "iam:permissions:revokeRoleFromAgencyOnDomain",
13        "iam:permissions:revokeRoleFromAgency",
14        "iam:permissions:grantRoleToAgencyOnDomain",
15        "iam:permissions:grantRoleToAgencyOnProject",
16        "iam:permissions:grantRoleToAgency",
17        "iam:permissions:listRolesForAgencyOnDomain",
18        "iam:permissions:listRolesForAgencyOnProject",
19        "iam:permissions:checkRoleForAgencyOnDomain",
20        "iam:permissions:checkRoleForAgencyOnProject",
21        "iam:permissions:listRolesForAgency",
22        "iam:permissions:checkRoleForAgency",
23        "iam:roles:listRoles"
24      ],
25      "Effect": "Allow"
26    }
27  ]
28 }
```

策略描述: 请输入策略描述 (可选)

作用范围: --

确定 取消

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "iam:agencies:createAgency",
        "iam:agencies:listAgencies",
        "iam:agencies:getAgency",
        "iam:agencies:deleteAgency",
        "iam:agencies:updateAgency",
        "iam:permissions:revokeRoleFromAgencyOnProject",
        "iam:permissions:revokeRoleFromAgencyOnDomain",
        "iam:permissions:revokeRoleFromAgency",
        "iam:permissions:grantRoleToAgencyOnDomain",
        "iam:permissions:grantRoleToAgencyOnProject",
        "iam:permissions:grantRoleToAgency",
        "iam:permissions:listRolesForAgencyOnDomain",
        "iam:permissions:listRolesForAgencyOnProject",
        "iam:permissions:checkRoleForAgencyOnDomain",
        "iam:permissions:checkRoleForAgencyOnProject",
        "iam:permissions:listRolesForAgency",
        "iam:permissions:checkRoleForAgency",
        "iam:roles:listRoles"
      ],
      "Effect": "Allow"
    }
  ]
}
```

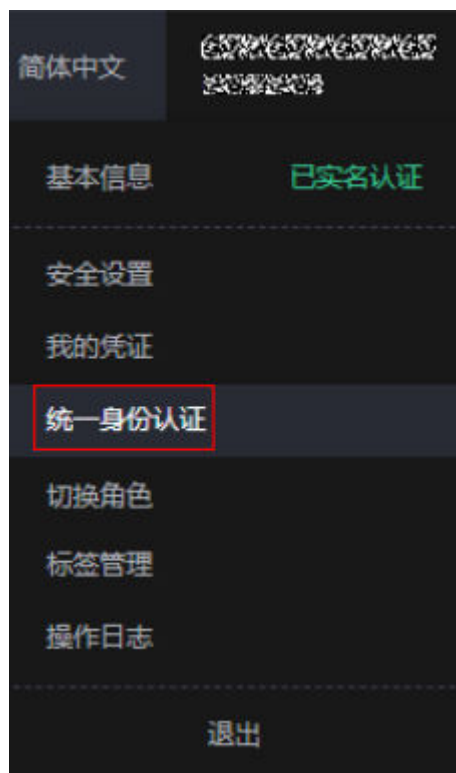
步骤5 界面无报错，则成功创建IAM Agency Management FullAcces权限。

----结束

## 给 rf\_admin\_trust 委托添加 IAM Agency Management FullAcces 策略

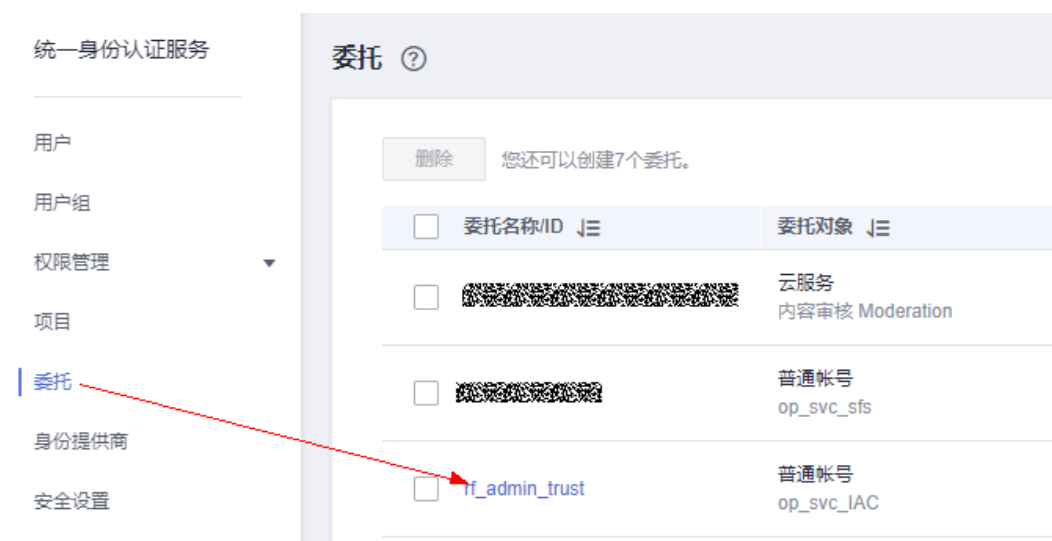
步骤1 打开“统一身份认证”菜单。

图 3-14 统一身份认证菜单



步骤2 进入“委托”菜单，选择rf\_admin\_trust委托。

图 3-15 委托列表



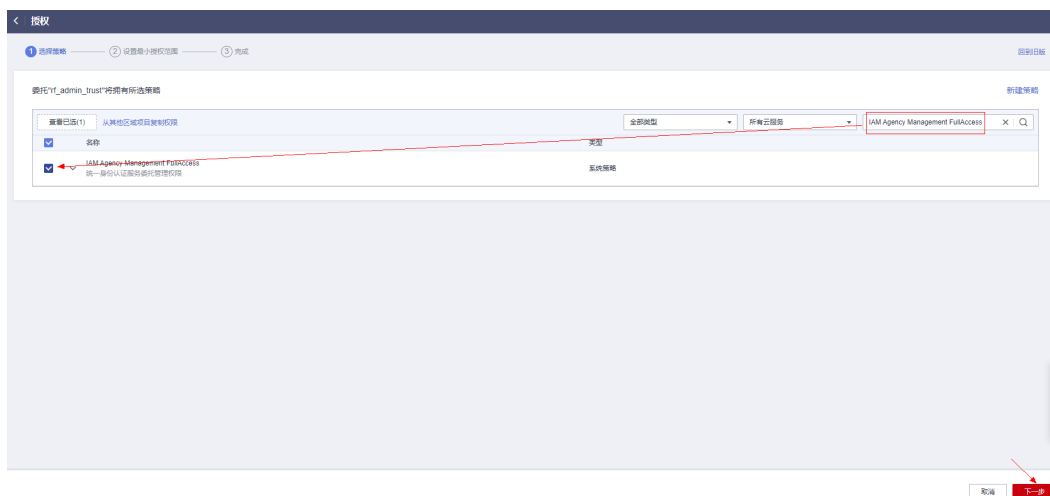
**步骤3** 进入“授权记录”菜单，单击“授权”按钮。

图 3-16 授权记录



**步骤4** 在搜索框输入IAM Agency Management FullAcces，勾选过滤出来的记录，单击下一步，并确认完成权限的配置。

图 3-17 配置 IAM Agency Management FullAcces 策略



**步骤5** 配置好后的情况：rf\_admin\_trust委托拥有Tenant Administrator和IAM Agency Management FullAccess权限。

图 3-18 授权记录列表



----结束

## 3.2 快速部署

本章节主要帮助用户快速部署“语音识别”解决方案。

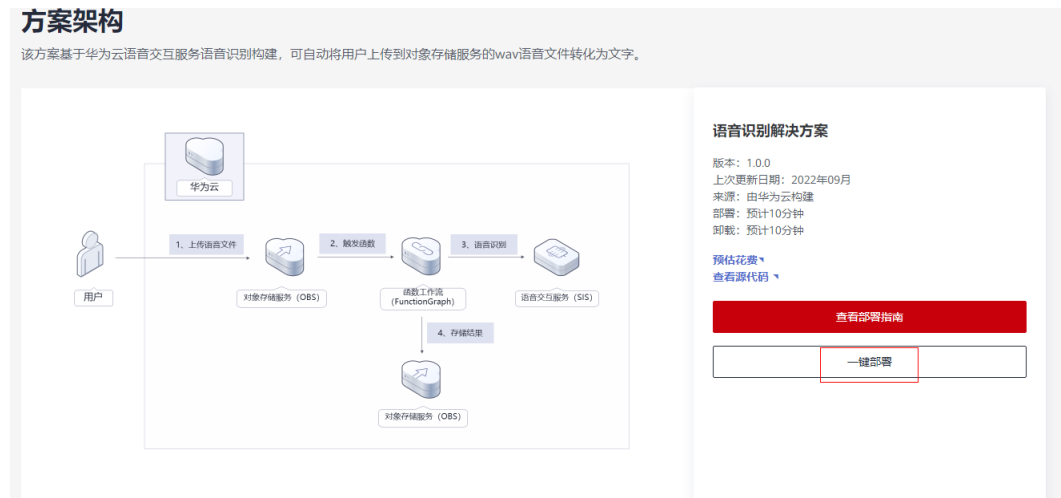
表 3-1 参数填写说明

参数名称	类型	是否必填	参数解释	默认值
function_name	String	是	函数名称，用于定义创建函数及其他资源前缀，不支持重名。取值范围：2-60个字符，可包含字母、数字、下划线和中划线，以大/小写字母开头，以字母或数字结尾。	speech-recognition-demo
speech_bucket_name	String	是	OBS桶名称，全局唯一，用于上传wav语音文件。取值范围：3~63个字符，支持小写字母、数字、中划线(-)、英文句号(.)。	
result_bucket_name	String	是	OBS桶名称，全局唯一，用于存放语音识别结果。取值范围：3~63个字符，支持小写字母、数字、中划线(-)、英文句号(.)。	

**步骤1** 登录[华为云解决方案实践](#)，选择“语音识别解决方案”，单击“一键部署”，跳转至解决方案创建堆栈界面。

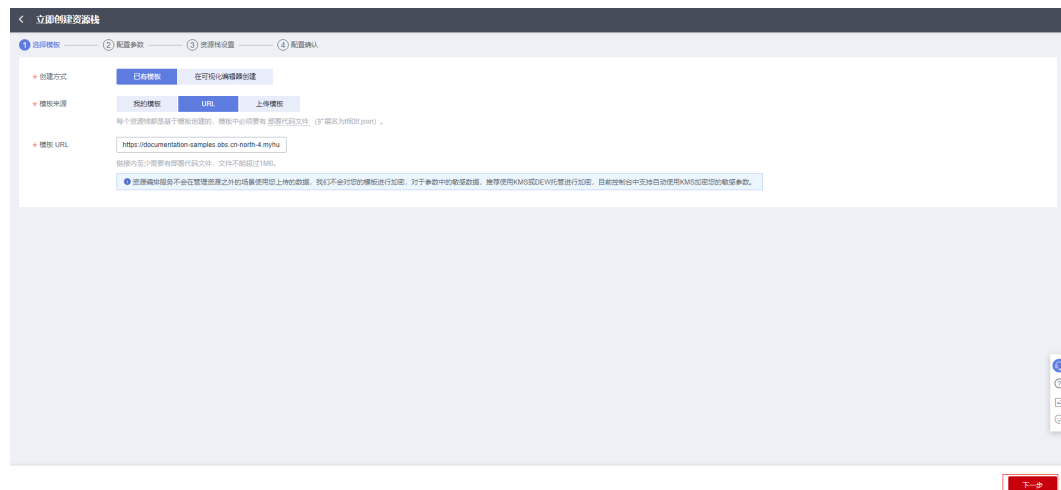


图 3-19 选择模板



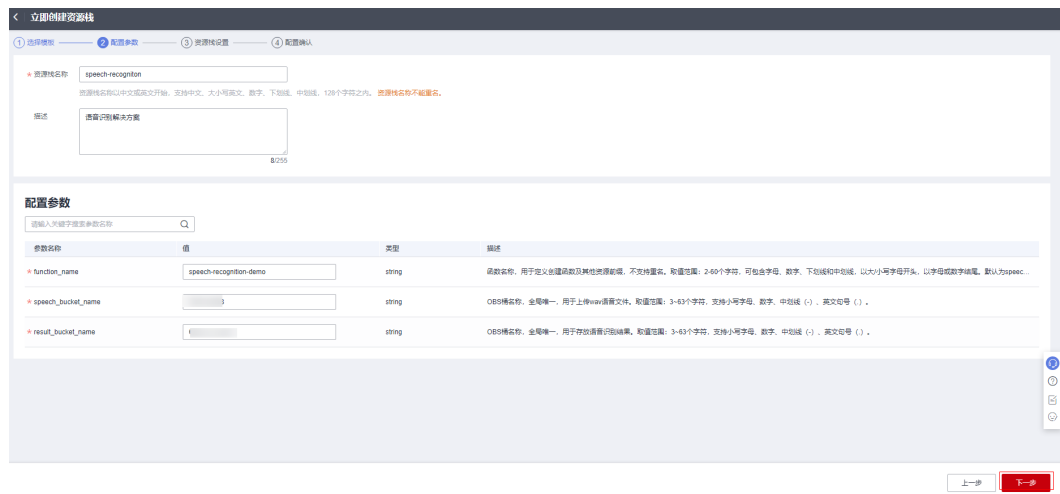
步骤2 在选择模板界面中，单击“下一步”。

图 3-20 选择模板



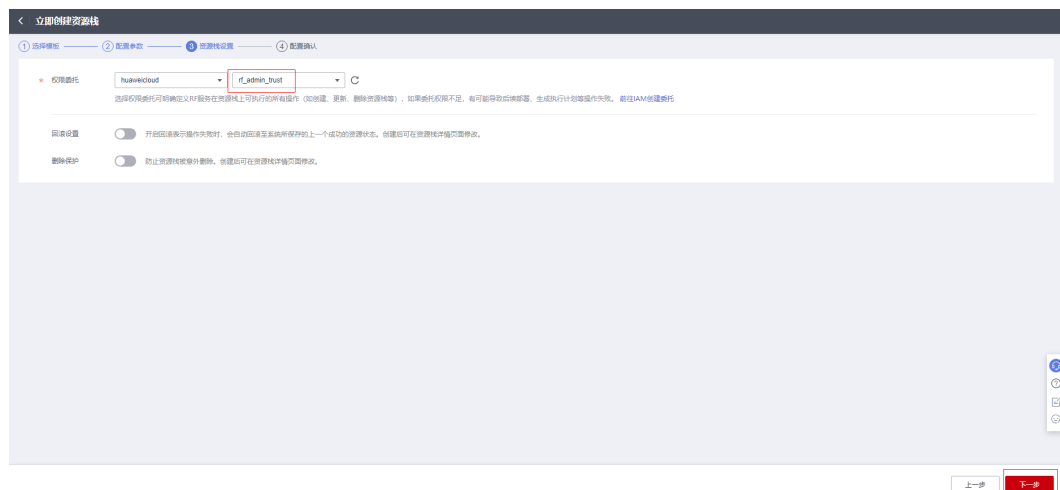
步骤3 在配置参数界面中，自定义填写堆栈名称，根据填写表3-1配置参数信息，单击“下一步”。

图 3-21 配置参数



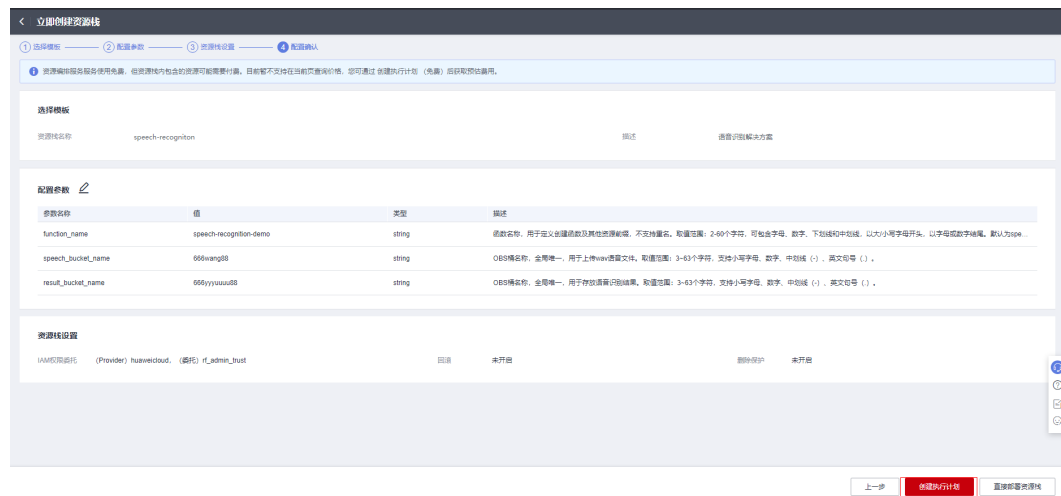
步骤4 在资源栈设置页面中，权限委托选择“rf\_admin\_trust”，单击“下一步”。

图 3-22 资源栈设置



步骤5 在配置确认页面中，单击“创建执行计划”。

图 3-23 配置确认



**步骤6** 在弹出的创建执行计划框中，自定义填写执行计划名称，单击“确定”。

图 3-24 创建执行计划



**步骤7** 等待执行计划状态为“创建成功，待部署”后，单击“部署”，并且在弹出的执行计划确认框中单击“执行”。

图 3-25 执行计划

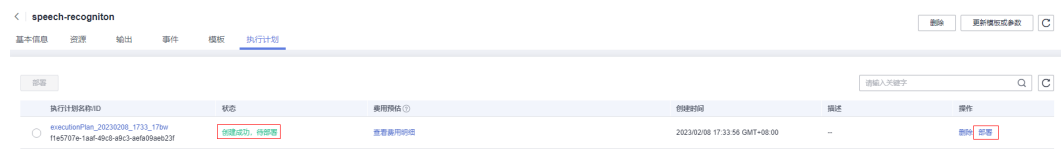


图 3-26 执行计划确认



**步骤8** 等待解决方案自动部署。部署成功后, 单击“事件”, 回显结果如下:

图 3-27 资源创建成功

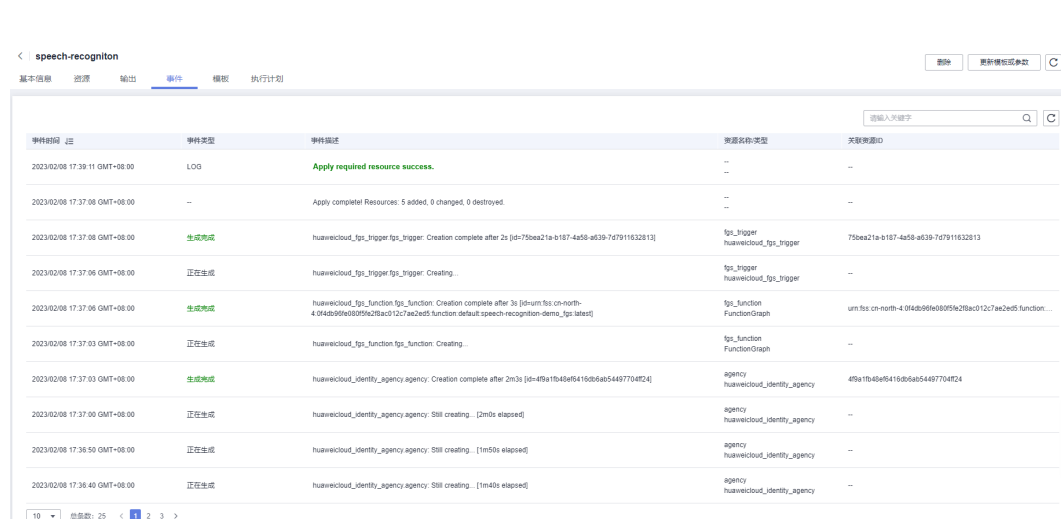
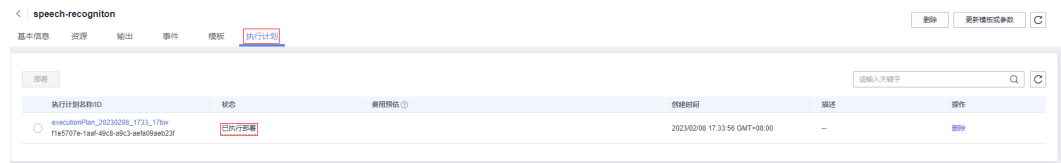


图 3-28 执行计划完成



----结束

### 3.3 开始使用

步骤1 登录华为云对象存储服务控制台，查看自动创建的OBS桶列表。

图 3-29 对象存储服务控制台



步骤2 选择用于上传和存储语音文件的桶“wwwy-1”（实际桶名称以部署指定参数为准），上传wav语音文件。

图 3-30 上传语音文件



步骤3 选择用于存放结果的“wwwy-2”桶（实际桶名称以部署指定参数为准），即可自动获取该语音文件的识别结果，以语音文件名称为前缀的JSON文件存储。

图 3-31 获取识别结果文件



----结束

## 3.4 快速卸载

### 删除委托

在进行一键卸载之前，需要先删除该方案创建的委托。

- 步骤1** 登录华为云**统一身份认证服务**控制台，在“委托”菜单栏中找到以该方案OBS桶名称为前缀的委托，单击“删除”。

图 3-32 删除委托



- 步骤2** 在弹出的删除确认框中单击“是”进行删除。

图 3-33 确认删除



----结束

## 一键卸载

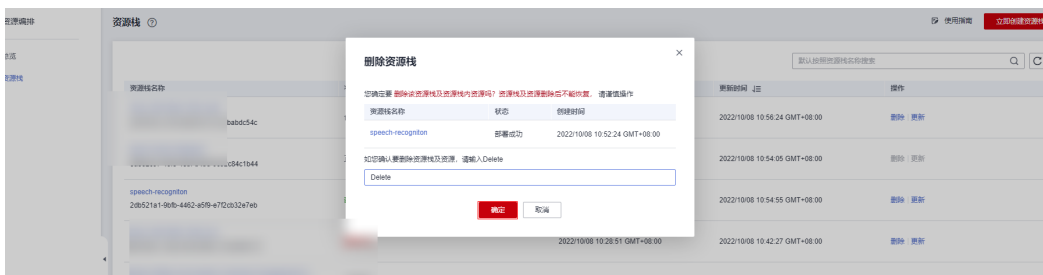
步骤1 解决方案部署成功后，单击该方案堆栈后的“删除”。

图 3-34 一键卸载



步骤2 在弹出的删除堆栈确认框中，输入“Delete”，单击“确定”，即可卸载解决方案。

图 3-35 删除堆栈确认



----结束

# 4 附录

## 名词解释

基本概念、云服务简介、专有名词解释

- 对象存储服务OBS：对象存储服务（Object Storage Service，OBS）是一个基于对象的海量存储服务，为客户提供海量、安全、高可靠、低成本的数据存储能力。
- 函数工作流FunctionGraph：FunctionGraph是一项基于事件驱动的函数托管计算服务。使用FunctionGraph函数，只需编写业务函数代码并设置运行的条件，无需配置和管理服务器等基础设施，函数以弹性、免运维、高可靠的方式运行。此外，按函数实际执行资源计费，不执行不产生费用。
- 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。例如用户通过语音识别功能，将口述音频或者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。



# 5 修订记录

发布日期	修订记录
2022-09-30	第一次正式发布。
2023-02-28	修订实施步骤。