

智能语音V2.X 语音合成SDK使用文档

功能介绍

调用限制

服务地址

交互流程

0.鉴权

1.start and confirm(设置请求参数)

2.synthesize audio data

3.stop and complete

服务状态码

SDK

Java SDK 2.0

C++ SDK 2.0

由SSML控制合成效果

功能介绍

语音合成提供将输入文本合成为语音二进制数据的功能。

- 支持输出pcm、wav、mp3编码格式数据；
- 支持设置语速，语调，音量。
- 支持设置男声/女声；

声音说明：

名称	voice参数值	类型	适用场景	支持语言	支持采样率 (Hz)
小云	xiaoyun	标准女声	通用场景	支持中文及中英文混合	8K/16K
小刚	xiaogang	标准男声	通用场景	支持中文及中英文混合	8K/16K

更多发音人请根据模型设置。

调用限制

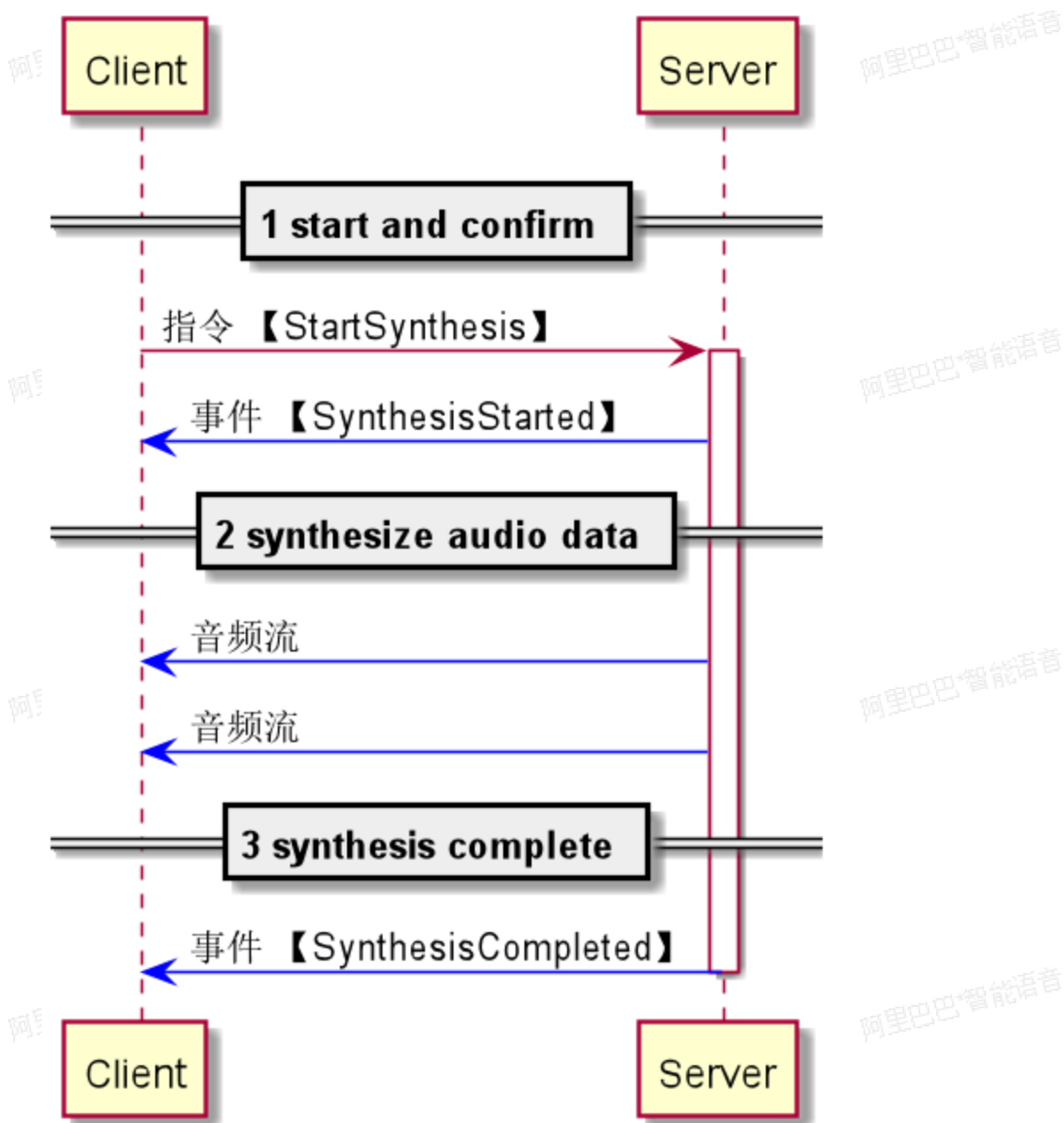
- 传入文本必须采用UTF-8编码；
- 传入文本不能超过300个字符。

服务地址

语音合成的WebSocket服务地址为：`ws://gateway所在IP:8101/ws/v1`。gateway默认的端口号为8101，如果部署时修改过gateway的端口号，请使用修改后的端口号。

交互流程

****说明：****交互流程图为Java SDK、C++ SDK、iOS SDK、Android SDK的交互流程，不包含RESTful API的交互流程。



说明：服务端的响应除了音频流之外，都会在返回信息的header包含task_id参数，用于表示本次识别任务的ID，请记录下这个值，如果发生错误，请将task_id和错误信息提交到工单或者语音接口人。

0.鉴权

客户端在与服务端建立WebSocket链接的时候，需要使用token进行鉴权。在专有云下，token的值为`default`。

1.start and confirm(设置请求参数)

客户端发起请求，服务端确认请求有效。其中在请求消息中需要进行参数设置，各参数由SDK中SpeechSynthesizer对象的相关set方法设置，各参数含义如下：

参数	类型	是否必须	说明
appkey	String	是	业务方或者业务场景的标记， 专有云默认为default
text	String	是	待合成文本， 文本内容必须采用UTF-8编码， 长度不超过300个字符 (英文字母之间需要添加空格)
voice	String	否	发音人， 默认是xiaoyun
format	String	否	音频编码格式， 默认是pcm。 支持的格式：pcm、wav、mp3
sample_rate	Integer	否	音频采样率， 默认是16000
volume	Integer	否	音量，范围是0~100， 默认50
speech_rate	Integer	否	语速， 范围是-500~500， 默认是0
pitch_rate	Integer	否	语调， 范围是-500~500， 可选，默认是0

2.synthesize audio data

服务端开始返回合成的语音二进制数据，SDK接收并处理二进制数

3.stop and complete

语音合成完毕，服务端发送合成完毕事件通知，示例如下：

```
{
  "header": {
    "message_id": "05450bf69c53413f8d88aed1ee600e93",
    "task_id": "640bc797bb684bd69601856513079df5",
    "namespace": "SpeechSynthesizer",
    "name": "SynthesisCompleted",
    "status": 20000000,
    "status_message": "GATEWAY|SUCCESS|Success."
  }
}
```

服务状态码

在服务的每一次响应中，都包含status字段，即服务状态码，状态码各种取值含义如下：

通用错误：

错误码	原因	解决办法
40000001	身份认证失败	检查使用的令牌是否正确，是否过期
40000002	无效的消息	检查发送的消息是否符合要求
40000003	令牌过期或无效的参数	首先检查使用的令牌是否过期，然后检查参数值设置是否合理
40000004	空闲超时	确认是否长时间没有发送数据掉服务端
40000005	请求数量过多	检查是否超过了并发连接数或者每秒钟请求数
40000000	默认的客户端错误码	查看错误消息或提交工单
50000000	默认的服务端错误	如果偶现可以忽略，重复出现请提交工单
50000001	内部GRPC调用错误	如果偶现可以忽略，重复出现请提交工单

网关错误：

错误码	原因	解决办法
40010001	不支持的接口	使用了不支持的接口， 如果使用SDK请提交工单
40010002	不支持的指令	使用了不支持的指令， 如果使用SDK请提交工单
40010003	无效的指令	指令格式错误， 如果使用SDK请提交工单
40010004	客户端提前断开连接	检查是否在请求正常完成之前关闭了连接
40010005	任务状态错误	发送了当前任务状态不能处理的指令

Meta错误：

错误码	原因	解决办法
40020105	应用不存在	检查应用appKey是否正确， 是否与令牌归属同一个账号

TTS错误：

错误码	原因	解决办法
41020001	参数错误	检查是否传递了正确的参数
51020001	TTS服务端错误	如果偶现可以忽略， 重复出现请提交工单

SDK

专有云只提供了Java和C++SDK，如果您需要使用移动端的Android或iOS SDK，请参考公共云使用方式：

- Android: https://help.aliyun.com/document_detail/84706.html
- iOS: https://help.aliyun.com/document_detail/84626.html

Java SDK 2.0

1. 下载安装

语音合成的目录为nls-example-tts，Demo依赖的jar包存放在lib目录，请参考pom文件添加的jar包依赖。

2. 关键接口

- NlsClient：语音处理client，相当于所有语音相关处理类的factory，全局创建一个实例即可。线程安全。
- SpeechSynthesizer：语音合成处理类，设置请求参数，发送请求。非线程安全。
- SpeechSynthesizerListener：语音合成监听类，监听返回结果。非线程安全。有如下两个抽象方法需要实现：

```
/**
 * 接收语音合成二进制数据
 */
abstract public void onMessage(ByteBuffer message);

/**
 * 语音合成结束事件通知
 *
 * @param response
 */
abstract public void onComplete(SpeechSynthesizerResponse response);
```

更多介绍参见API文档链接: [Java API接口说明](#)

3. SDK 调用注意事项

1. NlsClient对象创建一次可以重复使用，每次创建消耗性能。NlsClient使用了netty的框架，创建时比较消耗时间和资源，但创建之后可以重复利用。建议调用程序将NlsClient的创建和关闭与程序本身的生命周期结合。
2. SpeechSynthesizer对象不能重复使用，一个语音合成任务对应一个SpeechSynthesizer对象。例如有N个文本需要语音合成，则要进行N次语音合成任务，创建N个SpeechSynthesizer对象。
3. 实现的SpeechSynthesizerListener对象和SpeechSynthesizer对象是一一对应的，不能将一个SpeechSynthesizerListener对象设置到多个SpeechSynthesizer对象中，否则不能区分是哪个语音

合成任务。

4.Demo

Demo文件在SDK目录下的nls-example-tts子目录中：

文件名	路径	说明
SpeechSynthesizerDemo.java	src/main/java/com/alibaba/nls/client	语音合成单线程Demo
SpeechSynthesizerLongTextDemo.java	src/main/java/com/alibaba/nls/client	语音合成展示处理文本超过300字符的Demo
SpeechSynthesizerMultiThreadDemo.java	src/main/java/com/alibaba/nls/client	语音合成多线程Demo

所有语音合成参数设置方法：

```
//创建实例, 建立连接
SpeechSynthesizer synthesizer = new SpeechSynthesizer(client, getSynthesizerListener());
synthesizer.setAppKey(appKey);
//设置用于语音合成的文本
synthesizer.setText("北京明天天气怎么样啊");
// 设置发音人, 可选参数, 默认"xiaoyun"
synthesizer.setVoice("xiaoyun");
//设置返回音频的编码格式, 可选参数, 默认PCM
synthesizer.setFormat(OutputFormatEnum.WAV);
//设置返回音频的采样率, 可选参数, 默认16KHz
synthesizer.setSampleRate(SampleRateEnum.SAMPLE_RATE_16K);
// 设置音量, 可选参数, 默认50, 范围是0~100
synthesizer.setVolume(50);
// 设置语速, 可选参数, 默认0, 范围是-500~500
synthesizer.setSpeechRate(0);
// 设置语调, 可选参数, 默认0, 范围是-500~500
synthesizer.setPitchRate(0);
```

说明1： 专有云下，appkey为default，token为default，url为ws://gateway所在IP:8101/ws/v1。

说明2（仅限专有云2.5及以上版本）： Demo默认使用的是静态填写服务器地址的方式，使用VIPServer方式，SDK可动态获取存活的服务器地址，然后直连该服务器进行交互。使用步骤如下（可参考nls-example-vipclient目录下的NlsVipServerClientDemo）：

- 1.登录到APES页面获取gateway域名。
- 2.项目中引入 nls-vipserver-client 此库外网没有，需从代码示例nls-example-vipclient/lib中找到该库极其依赖自行导入项目
- 3.修改代码 NlsVipServerClient 继承自NlsClient，可以作为NlsClient初始化SpeechSynthesizer等类。SpeechSynthesizer创建成功后使用方式不变。

NlsVipServerClient初始化方式如下：

```
String vipServer = "您的APES所在服务IP";
String vipKey = "default.gateway.vipserver";
String token = "default";

NlsVipServerClient client = new NlsVipServerClient(vipServer, vipKey, token);
SpeechSynthesizer synthesizer = new SpeechSynthesizer(client, getSynthesizerListener());
```

登录到APES页面，上面代码中的三个参数vipServer， vipKey分别对应下面截图中的各个部分：



C++ SDK 2.0

1.下载安装

C++ SDK的压缩文件包含以下几个部分：

- CMakeLists.txt demo工程的CMakeList文件。
- build 编译目录。
- demo 包含demo.cpp，各语音服务配置文件。各文件描述见下表：

文件名	描述
sdkDemo.cpp	windows专用，默认为一句话识别功能demo，如需可自行替换成其它功能(编码格式：UTF-8代签名)
speechRecognizerDemo.cpp	一句话识别demo
speechSynthesizerDemo.cpp	语音合成demo
speechTranscriberDemo.cpp	实时音频流识别demo
speechRecognizerSyncDemo.cpp	一句话同步识别demo
speechTranscriberSyncDemo.cpp	实时音频流同步识别demo
testX.wav	测试音频

- include 包含sdk头文件，以及部分第三方头文件。各文件描述见下表

文件名	描述
openssl	openssl
pthread	pthread线（windows下使用）
uuid	uuid(linux下使用)
opus	opus
jsoncpp	jsoncpp
curl	nls SDK并不依赖curl，仅用于demo中，用以获取token)
nlsCommonSdknls	nls sdk并不依赖nlsCommonSdk，仅用于demo中，用以获取token
nlsClient.h	SDK实例
nlsEvent.h	回调事件说明
speechRecognizerRequest.h	一句哈识别
speechSynthesizerRequest.h	语音合成

speechTranscriberRequest.h	实时音频流识别
speechRecognizerSyncRequest.h	一句话同步识别
speechTranscriberSyncRequest.h	实时语音同步识别
iNlsRequest.h	Request基础

- lib 包含sdk, 以及第三方依赖库。linux.tar.gz为linux平台lib压缩包。windows.zip为windows平台lib压缩包。其中根据平台不同, 可以选择[linux版本libnlsCppSdk.so](#)(glibc2.5及以上, Gcc4, Gcc5), windows(x86/x64)版本nlsCppSdk.dll (VS2013、VS2015)。
- readme.txt SDK说明。
- release.log 版本说明。
- version 版本号。
- [build.sh](#) demo编译脚本。

2.依赖库

SDK 依赖 openssl(l-1.0.2j), opus(1.2.1), jsoncpp(0.10.6), uuid(1.0.3), pthread(2.9.1)。

依赖库放置在 path/to/sdk/lib 下。

注意:

path/to/sdk/lib/linux/uuid仅在linux下使用。path/to/sdk/lib/windows/1x.0/pthread仅在windows下使用。

3.编译运行

运行编译脚本build.sh:

1. 请确认本地系统以安装Cmake, 最低版本3.1、Glibc 2.5、Gcc 4.1.2及以上
2. `cd path/to/sdk/lib`
3. `tar -zxvpf linux.tar.gz`
4. `cd path/to/sdk`
5. 执行`./build.sh` 编译demo
6. 编译完毕, 进入path/to/sdk/demo目录。可以看见以生成三个demo可执行程序: srDemo(一句话异步识别)、srSyncDemo(一句话同步识别)、stDemo(实时音频异步识别)、stSyncDemo(实时音频同步识别)、syDemo(语音合成)。
7. 执行`./syDemo service-url`

如果不支持cmake, 可尝试手动编译:

- 1: `cd path/to/sdk/lib`
- 2: `tar -zxvpf linux.tar.gz`
- 3: `cd path/to/sdk/demo`

```

4: g++ -o syDemo speechSynthesizerDemo.cpp -I../include -L../lib/linux -lnlsC
ppSdk -lnlsCommonSdk -lcurl -lssl -lcrypto -lopus -lpthread -luuid -ljsoncpp
-D_GLIBCXX_USE_CXX11_ABI=0
5: export LD_LIBRARY_PATH=path/to/sdk/lib/linux/
6: ./syDemo service-url

```

Windows平台需要用户自己搭建工程。

注意：

1. linux环境下，运行环境最低要求：Glibc 2.5及以上， Gcc4、Gcc5
2. windows下，目前支持VS2013， VS2015

4.关键接口

- NlsClient：语音处理Client，相当于所有语音相关处理类的Factory，全局创建一个实例即可。
- SpeechSynthesizerRequest：语音合成请求对象。
- SpeechSynthesizerCallback：回调事件函数集合的对象，合成结果、异常等回调的统一入口。
- NlsEvent：回调事件对象，用户可以从其中获取Request状态码、云端返回结果、失败信息等。

更多介绍参见api文档链接:[C++ API接口说明](#)

5.C++ SDK自定义错误码

错误码	错误描述	解决方案
10000001	SSL: couldn't create a!	建议重试
10000002	openssl官方错误描述	根据描述提示处理之后， 建议重试
10000003	系统错误描述	根据系统错误描述提示处理
10000004	URL: The url is empty.	检查是否设置 云端URL地址
10000005	URL: Could not parse WebSocket url	检查是否正确设置 云端URL地址
10000006	MODE: unsupport mode.	检查时都正确设置了语音功能模 式

10000007	JSON: Json parse failed.	服务端发送错误响应内容， 请提供task_id， 并反馈给阿里云
10000008	WEBSOCKET: unkown head type.	服务端发送错误WebSocket类型， 请提供task_id， 并反馈给阿里云
10000009	HTTP: connect failed.	与云端连接失败， 请检查网络， 在重试
HTTP协议官方状态码	HTTP: Got bad status.	根据HTTP协议官方描述提示处理
系统错误码	IP: ip address is not valid.	根据系统错误描述提示处理
系统错误码	ENCODE: convert to utf8 error.	根据系统错误描述提示处理
10000010	please check if the memory is enough	内存不足. 请检查本地机器内存
10000011	Please check the order of execution	接口调用顺序错误 (接收到Failed/complete事件时， SDK内部会关闭连接。此时在调用send会上报错误。)
10000012	StartCommand/StopCommand and Send failed	参数错误。 请检查参数设置是否正确
10000013	The sent data is null or dataSize <= 0.	发送错误。 请检查发送参数是否正确
10000014	Start invoke failed.	start超时错误. 请调用stop， 释放资源，重新开始识别流程.
10000015	connect failed等	connect失败. 释放资源， 重新开始识别流程.

6.Demo

所有语音合成参数设置方法：

```

SpeechSynthesizerRequest* request = NlsClient::getInstance()->createSynthesizerRequest(callback);
// 设置服务端url, 必填参数
request->setUrl("ws://gateway所在IP:8101/ws/v1");
// 设置AppKey, 必填参数, 专有云为 default
request->setAppKey("default");
// 设置待合成文本, 必填参数. 文本内容必须为UTF-8编码
request->setText("北京的天气");
// 发音人, 默认"xiaoyun"
request->setVoice("xiaoyun");
// 音量, 范围是0~100, 可选参数, 默认50
request->setVolume(50);
// 音频编码格式, 可选参数, 默认是wav. 支持的格式pcm, wav, mp3
request->setFormat("wav");
// 音频采样率, 默认是16000
request->setSampleRate(SAMPLE_RATE);
// 语速, 范围是-500~500, 可选参数, 默认是0
request->setSpeechRate(0);
// 语调, 范围是-500~500, 可选参数, 默认是0
request->setPitchRate(0);
// 合成方法, 可选参数, 默认是0. 参数含义0:不带录音的参数合成; 1:带录音的拼接合成; 2:不带录音的拼接合成; 3:带录音的参数合成
request->setMethod(0);

```

完整示例, 详见SDK压缩包中的demo目录。

说明1: 专有云下, appkey为default, token为default, url为ws://gateway所在IP:8101/ws/v1。

说明2 (仅限专有云2.5及以上版本, Linux环境): Demo默认使用的是静态填写服务器地址的方式, 使用VIPServer方式, SDK可动态获取存活的服务器地址, 然后直连该服务器进行交互。使用步骤如下 (请参考vipServerDemo文件):

- 1.登录到APES页面获取gateway域名。
- 2.在调用setUrl()之前, 使用vipServerGetIp()获取语音服务ip地址即可。

```

/**
 * @brief 获取ip
 * @param vipServerDomain 地址服务器
 * @param port 地址服务器端口
 * @param targetDomain 语音服务域名
 * @param nlsServerUrl 语音服务目标url
 * @return 成功返回0, 失败返回-1

```

```

*/
static int vipServerGetIp(const std::string& vipServerDomain, const int vipServerPort, const std::string& targetDomain, std::string& nlsServerUrl);

```

Demo修改方式：

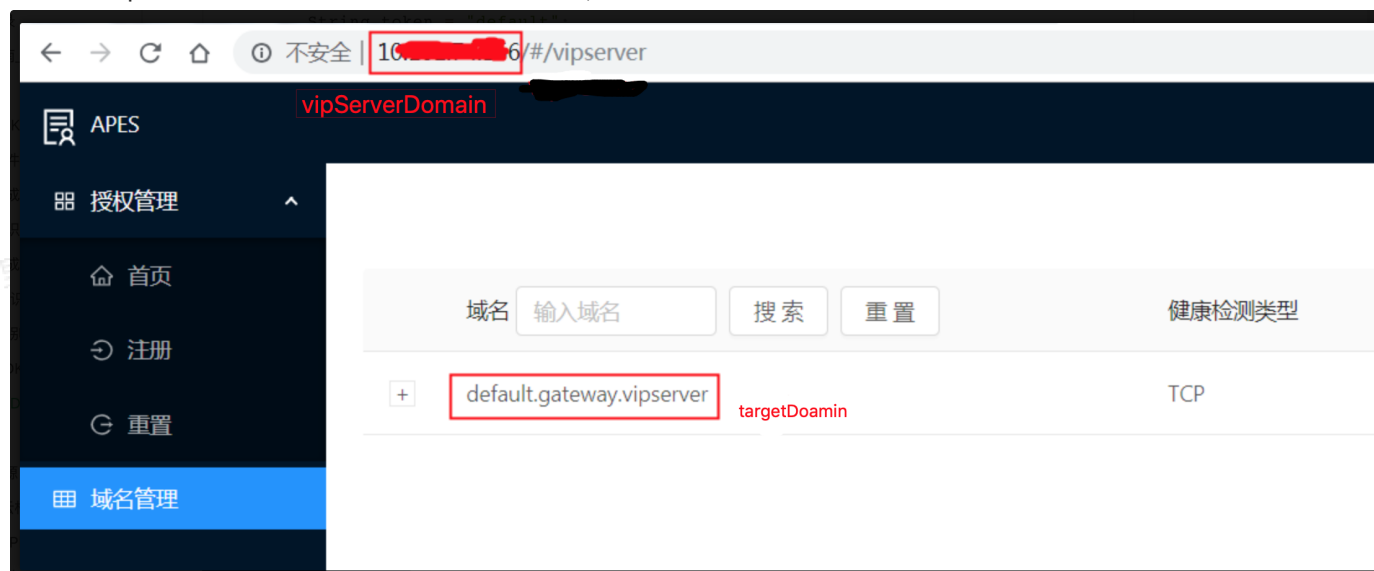
```

string nlsUrl;
// 获取ip与port
if ( -1 == NlsClient::getInstance()->vipServerGetIp(vipServerDomain, vipServerPort, targetDomain, nlsUrl)) {
    return NULL;
}

request->setUrl(nlsUrl.c_str());
.....

```

登录到APES页面，上面代码中的参数vipServerDomain, targetDomain分别对应下面截图中的红框部分。参数port可以设置为80。如果80不可用，请咨询部署人员。



由SSML控制合成效果

专有云支持的SSML标签与公共云支持的SSML标签相同，请直接阅读公共云SSML文档：

https://help.aliyun.com/document_detail/101645.html