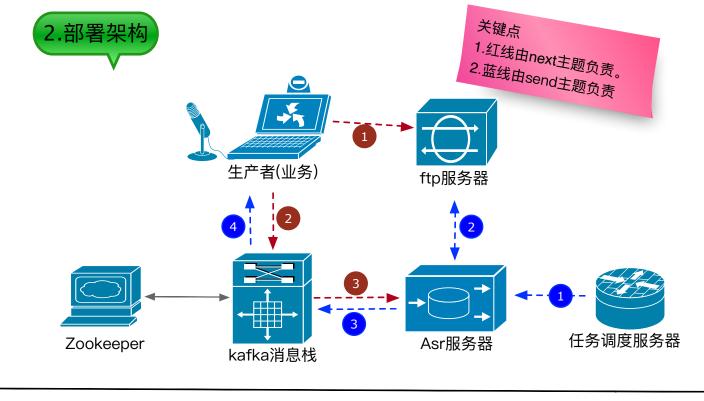
Asr项目测试准备资料

1.测试目的

- 1.对消息生产和消费流程重点关于并发性测试
- 2.通过大量的训练过的语句测试学习框架的准确性
- 3.通过未训练的语句测试目前模型库的问题,为下 一步自训练学习模型的设计积累经验
- 4.从功能和性能两方面验证目前分布式消息总线系统,为其他产品提供复用的可能性
 - 说,为其他产品提供复用的可能性 ————————————————————



3.关键点

- 1.生产者多线程并发消息,追查消费者日志是否有消息遗失?
- 2.生产者和消费者同时并发处理,追查反馈消息是否有遗失
- 3.测试训练过的样例语音测试文件,观察翻译的结果的错误率
- 5. 服务中生产者和消费者断电异常退出,查看消息是否遗失

4.录制未被训练过的语音文件,观察翻译结果的错误率

4.消息格式

1.对应ftp服务器存储目录,文件名按照日期和编号组合

"audioDir": "20191204/31011310200101000088_COVER_00002/audio",

"id":100,"

3.建立档案的时候,用户输入信息,用作自我训练用。

"textDir": "20191204/31011310200101000088_COVER_00002/text"

4.标记档案类型。当前为封面

"type":"COVER"

语音引擎

息4}

"message":{ "id":100, context:信息1|||信息2|||信息3|||信

接受消息格式

翻译的信息以||分割.id为发送的编号

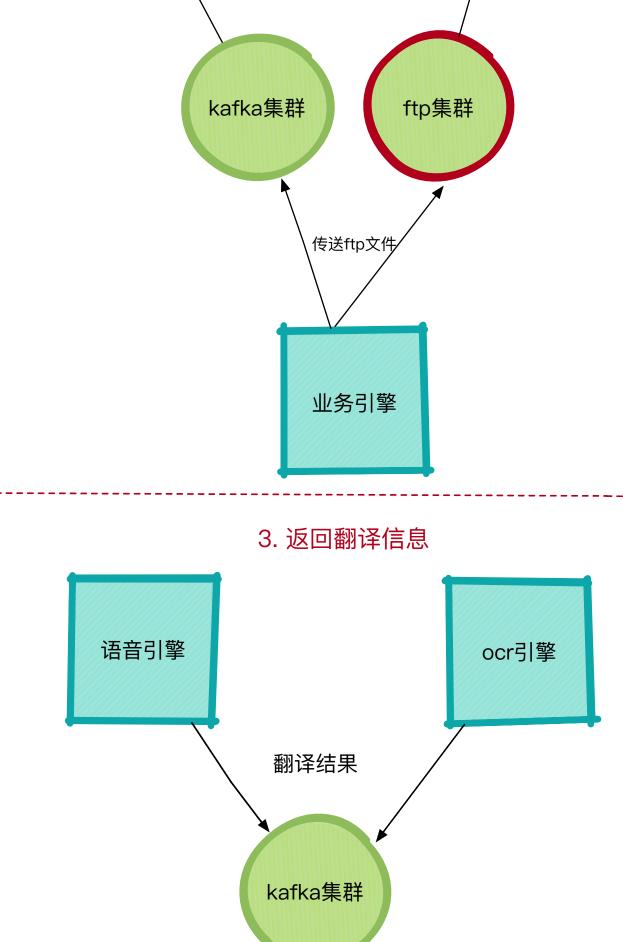
5.组件原理 A.消息类型

1. 注册 1).语音文件上传消息 2.)ocr图像上传信息 3.)翻译结果消息 B.语音和ocr图像消息 kafka集群 只携带上传文件的路径 信息,实际文件采用ftp 方式传输 注册 注册 注册 语音引擎 ocr引擎 业务引擎 2.上传声音和图像文件

接受消息

ocr引擎

接受消息



业务引擎

5.工具及配置

githup: https://github.com/
maxiaolu19710219/asr-test.git