**人脸服务总体设计**

**文档版本号 1.0.00**

**编 写 人： Closeli**

**编写时间：2018.04.20**

**部 门 名： Closeli**

**审 核 人：**

**审核时间：**

**修订历史**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **编号** | **修订内容描述** | **修订日期** | **修订后版本号** | **修订人** | **批准人** |
| 1 | 初稿 | 2018.04.20 | V1.0.00 |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 简介

本文档预期读者包括：项目经理，软件工程师，测试工程师等项目相关人员。

# 总体架构

## 架构V1.0



**说明：**

1. 对外交互
2. 能力平台通过API向负载中心获取人脸服务的地址等信息；
3. 能力平台与获取到的人脸服务进行业务交互。
4. 路由服务
5. 支持外部调用的API，用于返回负载最轻的人脸服务信息；
6. 支持通过Redis中各人脸服务上报的负载信息计算负载最轻的人脸服务。
7. 人脸服务
8. 支持外部调用的API；
9. 集成人脸检查算法库和人脸比对算法库；
10. 支持集群，并支持向Redis上报负载信息（当前业务负载，服务器配置信息等）；
11. 支持将人脸特征信息保存到数据库（MongDB）。

## 架构V2.0



# 路由服务

## 概述

1. 路由服务设计成一个公共服务，在开发不同的业务服务时可以共用。
2. 支持在配置文件中指定负载计算条件，如CPU使用率、内存使用率、业务负载量等。
3. 支持拒绝服务。

## 模块划分



**说明：**

1. Manager模块
2. ApiServer模块
3. LoadCalculator模块

（1）当前版本V1.0负载信息包括：CPU使用率、内存使用率。

1. Comm模块

## 时序图



**说明：**

## 模块设计

### 总体类图



**说明：**

## 监控项设计

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **序号** | **监控项** | **监控项名称** | **监控项类型** | **单位** | **上报频率** | **报警阈值** |
| 1 | 管理线程运行时间 | face\_server\_run\_time | Counter | 秒 | 30秒 | 不递增 |
| 2 | Redis访问耗时 | face\_server\_rds\_call\_time | Gauge | 毫秒 | 30秒 | >10000 |
| 3 | API接口访问次数 | face\_server\_api\_request\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 4 | 响应时间 | face\_server\_response\_time | Gauge | 毫秒 | 30秒 | >10000 |
| 5 | 鉴权失败次数 | face\_server\_authentication\_fail\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 6 | 接收未知请求次数 | face\_server\_api\_recv\_unkown\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 7 | API接口访问失败次数 | face\_server\_api\_fail\_counts | Gauge | 个 | 30秒 | 波动异常 |

# 人脸服务

## 概述

1. 支持过载保护（拒绝服务）。

## 模块划分



**说明：**

## 时序图

## 数据库设计

### 服务负载信息

服务负载信息由每个人脸服务（FaceServer）生成，并上报到统一的Redis数据库中；路由服务（RouteServer）从Redis上获取所有的服务负载信息，并计算负载值。

服务负载信息在Redis中以字符串形式存储，数据结构包括Key和Value。

Key必须由人脸服务设置生存时间，并在过期前更新生存时间，以保证服务不存在时负载信息能够自动删除。Key组成如下：

|  |
| --- |
| Key = "load.info.str.fs:" + [16进制服务IP] + "-" + [16进制服务端口] |

Value为服务负载信息，是由Json序列化后的字符串，具体内容如下：

|  |  |  |
| --- | --- | --- |
| **Key** | **Value** | |
| **Example** | [**Description**](file:///C:\Users\yyh7635\AppData\Local\Youdao\Dict\Application\7.5.0.0\resultui\dict\?keyword=description) |
| load.info.str.fs:ac1d12f8-1f9a | { |  |
| "SvrInfo":{ | 服务信息 |
| "Ip":"172.29.11.52", | 服务IP地址 |
| "Port":8090, | 服务端口 |
| } |  |
| "LoadInfo":{ | 负载信息 |
| "CpuUsage":30, | CPU使用率[0, 100] |
| "MemUsage":20, | 内存使用率[0, 100] |
| } |  |
| } |  |

### 人脸特征信息

## 模块设计

### 总体类图



**说明：**

### Manager模块

* + - 1. **类图**
      2. **类定义**
      3. **流程图**

## 监控项设计

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **序号** | **监控项** | **监控项名称** | **监控项类型** | **单位** | **上报频率** | **报警阈值** |
| 1 | 管理线程运行时间 | face\_server\_run\_time | Counter | 秒 | 30秒 | 不递增 |
| 2 | Redis访问耗时 | face\_server\_rds\_call\_time | Gauge | 毫秒 | 30秒 | >10000 |
| 3 | API接口访问次数 | face\_server\_api\_request\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 4 | 响应时间 | face\_server\_response\_time | Gauge | 毫秒 | 30秒 | >10000 |
| 5 | 鉴权失败次数 | face\_server\_authentication\_fail\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 6 | 接收未知请求次数 | face\_server\_api\_recv\_unkown\_counts | Gauge | 个 | 30秒 | 波动异常 |
| 7 | API接口访问失败次数 | face\_server\_api\_fail\_counts | Gauge | 个 | 30秒 | 波动异常 |