空调大数据管理系统设计说明书

1 引言

1.1编写目的

本文档包括格力《空调大数据管理系统》的总体设计与系统架构，程序模块划分与程序代码结构等，为该系统的开发人员，二次开发人员，系统维护人员提供参考

1.2背景

软件名称 : 格力空调大数据管理系统

项目提出者 : 格力电气

项目开发者 : 华中科技大学

系统开发环境 : eclipse CDT(linux)

大数据文件系统 : hadoop hdfs

1.3定义

1.4参考资料

多联机GPRS传输通讯协议.pdf

空调大数据管理系统协议（客户端部分）.pdf

2. 总体设计

2.1需求规定

2.1.1系统功能

该系统实现 ：

1. 接收来自不同gprs系统的数据，并分不同gprs系统将gprs数据按接收的先后顺序保存到hadoop hdfs 上，客户端可以根据不同gprs系统请求对应的数据;

2. 能够接收监听客户端请求，将监听客户端请求监听的gprs数据实时转发到对应的客户端。

3. 从hdfs 上取出数据发送给请求数据的解析客户端

2.1.2系统性能

2.1.2.1并发量

服务器最大能够同时连接的数据不少于 10000

2.2运行环境

2.2.1支持软件

rhel6 , boost1.54.0 , zlog1.2.12 , JDK1.8.0\_60 , g++4.8.4 , hadoop2.7.1 ,eclipse Mars Release(4.5.0) CDT8.6.0

2.2.2接口

该系统与gprs模块和客户端之间通过 TCP/IP 协议通信

接入服务器与转发服务器之间使用 unix 域套接字通信

3运行设计

3.1运行模块组合

系统分为四个模块，各个模块之间相互独立，分别为:接入服务器，转发服务器，上传服务，接口，

接入服务器与转发服务器之间通过 unix 套接字连接

接入服务器与上传服务通过本地文件通信，接入服务器将收到的数据存入本地文件，上传服务读取本地文件，并将其上传到 hdfs

接口直接从 hdfs 上读取数据，并将基发送到请求的解析客户端

4. 系统实现

4.1 系统总体结构图

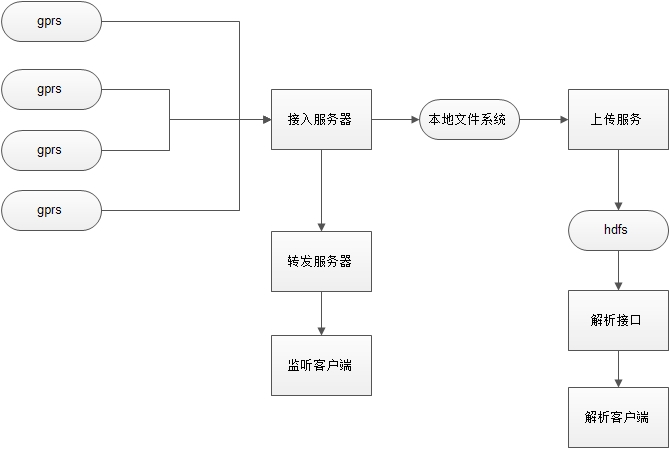


图 1 系统总体结构图

* 接入服务器

接入服务器实现两个功能，接收 gprs 数据，将接收到的gprs数据存入本地文件系统; 将监听客户端请求的gprs数据转发到对应的客户端。

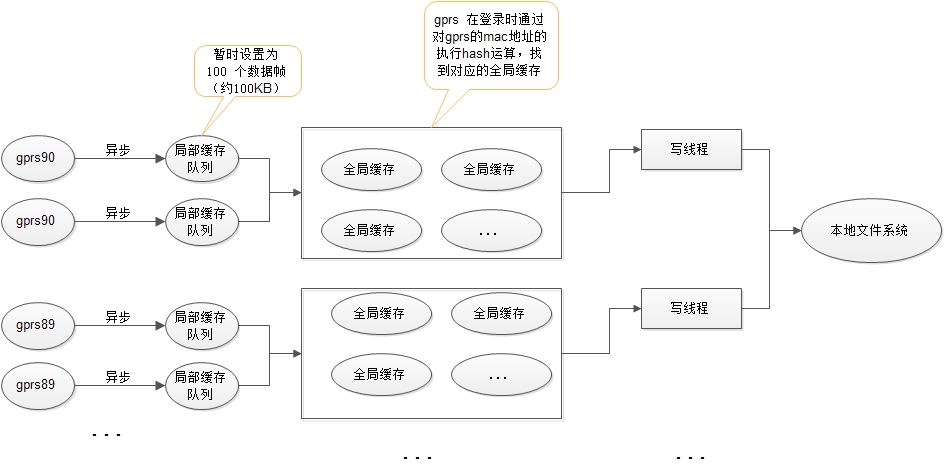
* 转发服务器

代理监听客户端的请求，与接入服务器进行交互，将监听客户端发送的请求转发给接入服务器并将接入服务器发送的数据转发给对应的客户端。

* 上传服务

负责将接入服务器存储到本地文件系统的文件上传到 hdfs。

4.2 接入服务器结构图

图 2 接入服务器程序结构图

接入服务器采用二级缓存，每个gprs连接都会都会分配一个队列，用于缓存收到的数据。这部分缓存称为局部缓存，局部缓存只与每个连接有关，只能在连接内部访问，其它部分不能访问，因此不用加锁。局部缓存的大小为100 个报文的大小。当缓存的数据达到缓存大小时，将其放到全局的缓存中。

全局缓存为各个连接与写文件线程共享，因此需要加锁以防冲突。全局缓存根据不同系统划分，每个系统有自己对应的全局缓存，每个全局缓存分为 100 个队列，当gprs 连接到达时，会根据源 mac 地址的hash 值从全局缓存中选择一个队列，当gprs局部缓存存满时，会将其数据放到选择的队列中。全局队列中的数据会根据内存自动扩充。

写文件线程负责将数据写入文件系统。写文件线程从全局队列中取出数据，并将数据写入本地文件系统。线程根据线程创建，每个系统对应一个写文件线程。

短信客户端负责与短信服务器通信。当监听客户端的请求的gprs模块没有登陆时，接入服务器将监听服务器的请求转交给短信客户端，则短信客户端将请求发送到短信服务器，请求打开对应的gprs。

4.3 转发服务器结构图

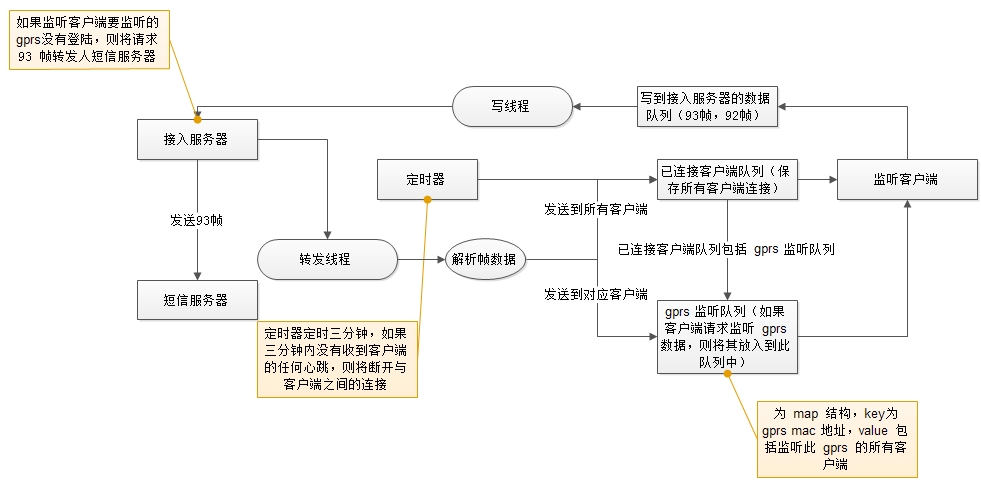
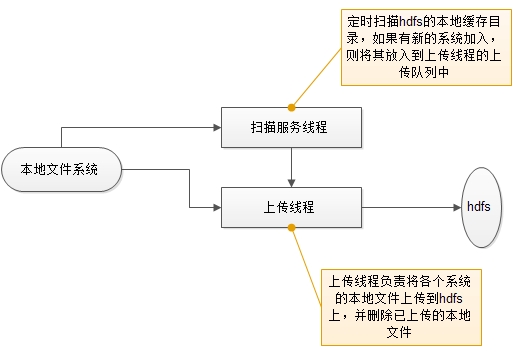


图 3 转发服务器结构图

转发服务器的数据结构主要分为两个 map 结构，一个用于保存所有已经连接的客户端，可用于转发数据到所有客户端。一个用于保存已连接的gprs的数据信息及监听此gprs的所有客户端，用于转发gprs数据到监听客户端。

4.4 上传服务结构图

图 4 上传服务程序结构图

上传服务分为两部分，分别为两个线程:

一个线程用于扫描接入服务器存放文件的路径，如果有新的文件则将其添加到需要上传的队列中。

一个线程用于读取每个系统下对应的文件，并将其上传到hdfs上。

5. 程序代码结构

5.1 接入服务器

* src/main.cpp : 程序运行主函数，包括读取配置文件，获取 mac 地址，初始化日志系统等，并启动短信客户端与gprs 侦听端口
* src/server/tsdb\_session.h tsdb\_session.cpp : 第个客户端对应一个 tsdb\_session 对象，该对象的任务包括读取gprs 发送的数据，将数据划分为完整的帧结构，并将其存入局部缓存，进而保存到全局缓存供写线程读取;将接入服务器需要发送给gprs的数据发送给gprs。
* src/server/tsdb\_server.h tsdb\_server.cpp : 服务器的主要部分，功能包括监听gprs连接，如果连接到来，则创建 tsdb\_session 处理gprs数据; 监听转发服务器连接，读取并转发转发服务器命令并将gprs数据转发给转发客户端; 保存已连接的gprs模块信息; 开启写线程从全局队列中取出数据并存储到本地文件系统; 如果转发服务器请求的gprs没有登陆，则将监听命令放入发送短信队列中; 启动定时器，如果两分钟后gprs 仍然没有登陆，则返回客户端 EE09 数据帧。
* src/server/MessageClient.h MessageClient.cpp : 负责与短信服务器，将发送队列中的监听命令发送给短信服务器。
* src/server/LOG.h LOG.cpp : 日志类，用于保存日志初始化信息。
* Src/server/ListWrapper.h ListWrapper.cpp : 对 list 的封装，实现多线程可以并发访问。
* src/server/HDFSReadWrite.h HDFSReadWrite.cpp : 负责将字节数据写入文件
* src/server/GetSysMac.h GetSysMac.cpp : 获取本机的 mac 地址
* src/server/Debug.h : 输出信息到日志文件
* src/server/data\_frame.h data\_frame.cpp : 保存协议的一帧数据，并提供操作协议帧的方法。
* src/server/BufferedFileName.h BufferedFileName.cpp : 获取准备写的文件名，并打开对应的流返回给调用者。
* src/UTIL.h UTIL.cpp : 程序中用到的工具函数

5.2 转发服务器

* src/ClientServer.h ClientServer.cpp : 接收客户端连接，并处理客户端监听与停止监听命令，如果要监听的gprs已经在监听，则直接将数据发送到监听客户端，如果gprs没有客户端在监听，则将请求转发给接入服务器。启动心跳线程，如果客户端3分钟没有发送心跳包则断开客户端的连接。接收接入服务器转发过来的数据，并将其转发到对应的客户端。
* src/ClientSession.h ClientSession.cpp : 接收监听客户端发送过来的数据，将交给 ClientServer 处理; 将数据发送到客户端。
* src/data\_frame.h data\_frame.cpp : 协议帧的一帧数据，并有操作协议帧的一系统方法。
* src/Debug.h : 输出日志信息
* src/GetSysMac.h GetSysMac.cpp : 获取本机 mac 地址
* src/ListWrapper.h ListWrapper.cpp : 对 list 的封装，实现多线程可以同时访问。
* src/LOG.h LOG.cpp : 保存日志初始化信息
* src/main.cpp : 主程序,包括读取配置文件，获取mac地址，初始化日志系统，开启监听客户端连接。
* src/UTIL.h UTIL.cpp : 程序中用到的工具函数

5.3 上传服务

* src/BufferedFileName.h BufferedFileName.cpp : 读取文件名配置文件，保存各个系统正在上传的文件名并保存到配置文件中。获取需要上传的文件流。
* src/config.h getConfig.cpp : 读取配置文件的方法
* src/Debug.h : 输出日志信息
* src/HdfsConnectionPool.h HdfsConnectionPool.cpp : 获取与hdfs 的连接
* src/server/HDFSReadWrite.h HDFSReadWrite.cpp : 负责将字节数据写入hdfs
* src/LOG.h LOG.cpp : 保存日志初始化信息
* src/main.cpp : 读取配置文件，初始化日志系统。
* src/UTIL.h UTIL.cpp : 程序中用到的工具函数
* src/uploadService.h uploadService.cpp : 定期扫描 /root/hadoop/ 文件夹，并将其中所有的文件夹添加到上传队列中; 扫描各个系统下的文件，如果找到可以上传的文件，则将其上传到 hdfs上。