车船税联网征收平台数据采集模块

目前,在总局的领导下，全国各地都在逐渐的部署公司研发的车船税连网征收系统，有超过6个地区已经部署了车船税连网征收系统并已经投入了运行。

现在，各地区为车船税联网征收系统提供的硬件设备以及数据库的构建模式（采用归档模式和非归档模式两种）各不相同，所以各地区车船税联网征收生产系统所受到的压力也有所不同，为了对系统中的信息进行利用并减少对生产系统的压力，建设备份库显得尤为重要，对于采用了归档模式建设的数据库系统可以采用数据库自带的备份机制进行数据备份形成备份库，对于采取非归档模式建设的数据库系统采用建设数据库过程（数据采集）的方式建设备份库，本文所描述的就是采用数据库过程（数据采集）的方式建设备份库，由于目前时间比较紧任务多，这里主要论述对联网征收主要数据信息的采集和备份，其他的信息日后会逐渐完善。

1. 数据采集模块架构

数据采集模块采用结构化的方式进行构建，具体如下图所示：

数据采集模块的实现的目标就是在定时器的管理下，及时、准确、快速的将联网征收平台的生产库的相关信息完整的采集备份到各地区车船税联网征收备份系统的备份库中，为后续的处理和分析减小对生产库压力并提供数据源。

1. 模块功能

本模块主要包括：数据及时性检查、数据完整性检查、数据抽取备份、日志处理、定时器建设等五项功能。

1、数据及时性检查

定时对采集到备份库中的数据进行检查，确认数据是否已经采集（备份）到备份库，如果数据已经采集（备份）到备份库则检查结束，否则检查生产库是否有数据，如果没有则检查结束，否则进行数据采集（备份）。

2、数据完整性检查

定时对采集到备份库中的数据（数据条数）与生产库中要采集（备份）的数据（条数）进行比对分析，如果一致说明本次采集（备份）正常完成，否则删除备份库中采集（备份）的数据从新进行数据采集（备份）。

3、数据采集备份

在定时器的控制下，对生产库中要备份的数据进行采集并备份到备份库中，对采集（备份）的结果记录到日志文件中。

4、日志文件处理

日志文件记载数据采集（备份）工作的所有状态，包括数据表名称、采集（备份）日期、完成标志、处理数据条数等。

5、定时器建设

定时器是实现自动数据采集（备份）的控制单元，根据采集（备份）的工作流程，控制和启动各个处理功能。

1. 采集（备份）方式

采集（备份）的方式包括两种，一是根据采集（备份）的条件进行增量采集（备份），再就是进行全量备份。

|  |  |  |  |
| --- | --- | --- | --- |
| 数据表 | 采集方式 | 控制方式 | 启动时间 |
| 代收代缴入库明细数据的采集 | 增量采集 | 时间段 | 每日24时 |
| 查询入参表的采集 | 增量采集 | 时间段 | 每日24时 |
| 确认入参表的采集 | 增量采集 | 时间段 | 每日24时 |
| 代收代缴欠税明细数据的采集 | 增量采集 | 时间段 | 每日24时 |
| 变更查询入参表的采集 | 增量采集 | 时间段 | 每日24时 |
| 变更确认入参表的采集 | 增量采集 | 时间段 | 每日24时 |

1、增量采集

对于第一次数据采集（备份）而言，所有的数据进行全量采集（保证历

史数据的完整性）。

后续的数据采集根据上表的采集方式进行增量数据采集（备份）。

1. 全量采集

对备份库中存在的要进行全量采集的表中的数据进行清除，然后把生产库中相应的表中的数据全部写入到备份库相应的表中。

1. 对于第一次数据采集（备份）而言，所有的数据进行全量采集要保证采用（备份）流程控制。

采集（备份）的流程由定时器进行控制，定时器根据每项工作的检查结果及完成情况来确定启动不同的任务执行。总体的控制流程如下：

1. 时间段控制

主要信息表的数据采集根据时间段来确认，对于代收代缴入库明细数据表、查询入参表、确认入参表、代收代缴欠税明细数据而言，时间段对应各表的“系统采集日期”，只要是“系统采集日期”在时间段范围内的数据全部进行采集并写入备份库中对应的数据表。

1. 备份库数据表的建设

在备份库中建设与生产库中相应的数据表，表名称与字段名称与生产库中的表名称和字段名称一致。