

用电信息采集前置性能提升关键技术研究与应用

赵俊¹, 李艳西¹, 何正民^{2*}, 齐明¹, 陈石东¹, 史科科³, 詹军², 赵亚茹²

¹国网湖南省电力有限公司信息通信分公司, 湖南 长沙

²北京科东电力控制系统有限责任公司, 北京

³南京新联软件有限责任公司, 江苏 南京

Email: 411799731@qq.com

收稿日期: 2021年1月6日; 录用日期: 2021年1月28日; 发布日期: 2021年2月8日

摘要

当前用电信息采集系统的采集前置部分在海量终端的接入和通信、海量高频数据的采集和入库等方面存在明显的技术瓶颈, 任务处理灵活性不高、并发处理能力不足、扩展性较弱, 难以满足新形势下各类业务日益增长的多元化需求。针对这些问题, 提出了一套解决采集前置性能瓶颈的方案, 涵盖通信任务分级管理、采集任务并行执行、高并发通信处理、规约动态扩展及适配、采集数据入库并行处理等多个方面。在国网某省级电力公司用电信息采集主站提升改造项目中的工程应用表明, 前置性能明显提升。

关键词

采集前置, 任务分级管理, 消息队列, 高并发处理, 规约动态适配

Research and Application of Key Technologies of Performance Improvement in Electric Energy Data Acquisition Front-End System

Jun Zhao¹, Yanxi Li¹, Zhengmin He^{2*}, Ming Qi¹, Shidong Chen¹, Keke Shi³, Jun Zhan², Yaru Zhao²

¹Information and Communication Branch of State Grid Hunan Electric Power Co., Ltd., Changsha Hunan

²Beijing Kedong Electric Power Control System Co., Ltd., Beijing

³Nanjing Xinlian Software Co., Ltd., Nanjing Jiangsu

*通讯作者。

文章引用: 赵俊, 李艳西, 何正民, 齐明, 陈石东, 史科科, 詹军, 赵亚茹. 用电信息采集前置性能提升关键技术研究与应用[J]. 智能电网, 2021, 11(1): 12-18. DOI: 10.12677/sg.2021.111002

Email: 411799731@qq.com

Received: Jan. 6th, 2021; accepted: Jan. 28th, 2021; published: Feb. 8th, 2021

Abstract

At present, there are obvious technical bottlenecks in the access and communication of massive terminals, and the collection and storage of massive high-frequency data in the acquisition front-end part of the current electric energy data acquisition system. The task processing flexibility is not high, the concurrent processing ability is insufficient, and the expansibility is weak, which is difficult to meet the diversified needs of various businesses under the new situation. In order to solve these problems, a set of solutions to the performance bottleneck of acquisition front-end system is proposed, which covers communication tasks classification management, data collection task parallel execution, high concurrent communication processing, protocol dynamic adaptation, data storage parallel processing and so on. The engineering application in the upgrading and reconstruction project of the master station of electric energy data acquisition system in a provincial company of the State Grid shows that the performance of acquisition front-end system is significantly improved.

Keywords

Acquisition Front-End System, Task Hierarchical Management, Message Queue, High Concurrency Processing, Protocol Dynamic Adaptation

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

电力用户用电信息采集系统(以下简称“采集系统”)是对电力用户的用电信息进行采集、处理和实时监控的系统[1]。采集系统结构复杂,包括前台应用、采集前置、通信信道、采集和计量设备等多个部分[2]。采集前置是采集系统的关键部分,主要承担采集系统主站与终端通信、任务管理、数据采集解析及入库等功能。

近年来,随着能源互联网[3][4]建设的持续推进,智慧能源、非侵入式负荷识别、电力现货市场交易等新兴业务[5][6][7]不断涌现,采集系统接入了更多类型和数量的设备[8][9],采集频率提高,交互数据量越来越庞大,呈爆发式增长,达到TB甚至PB等级[10]。当前采集系统的采集前置部分在海量终端的接入和通信、海量高频数据的采集和入库等方面存在明显的技术瓶颈,任务处理灵活性不高、并发处理能力不足、扩展性较弱,难以满足新形势下各类业务日益增长的多元化需求。因此,采集前置的性能亟待优化改善。

目前已有的一些提高采集前置性能的方法。文献[11]提出了基于Spark集群的多任务自适应的实时调度方法,可提高任务分配的合理性和高优先级任务的响应速度,其要求采集前置采用Spark平台架构;文献[12]采用Redis设计高并发应用系统,可提升信息系统的并发性,但其设计针对的是选课系统,并发数据量远低于采集系统,且Redis的核心功能为内存数据库,吞吐能力相对较低;文献[13]提出一种应用

Apache Kafka 分布式消息队列的方法，可提升海量数据的入库性能，该方法未考虑数据库类型和存储结构的影响；文献[14]基于面向对象和适配技术提出了通用型规约解析引擎模型，使其可以通过配置来实现对各种通信规约的适配应用，但系统模型相对比较简单，在面对庞大的采集系统业务时，在实践中需要采用更多的优化措施来应对复杂的业务逻辑。另外，影响采集前置性能的因素有很多，以上方法都只是从某一方面优化采集前置的性能，未从整体上考虑各方面存在的关联性。

本文首先分析了采集前置目前存在的问题，然后提出了一套解决采集前置性能瓶颈的方案，涵盖通信任务分级管理、采集任务并行执行、高并发通信处理、规约动态扩展及适配、采集数据入库并行处理等多个方面。最后在国网某省电力公司用电信息采集主站提升改造项目中进行了工程应用，取得了良好的效果。

2. 关键技术研究

采集前置性能涉及多个方面，影响因素也很多，本文从影响采集前置性能的关键环节进行研究处理，包括通信任务分级管理、采集任务并行执行、高并发通信处理、规约动态适配、采集数据入库并行处理等方面。

2.1. 通信任务分级管理

当前采集前置按接收顺序依次对所有下行请求指令进行报文封装处理，即按照“先到先得”的执行策略，当出现下行请求阻塞引起的指令排队等待情况时，重要指令下发的实时性得不到保证，从而影响某项相关业务的正常开展。

针对此问题，首先需要对通信任务进行分级管理。根据通信任务的紧急程度、重要性、用户体验等方面的不同，通信优先级分为高、中、低三个档次。基本的分级原则是：控制指令优先级最高，页面操作高于自动任务。为首先保证主站业务，外部系统操作发起的通信任务的优先级在每个分档中处于中下段。按表 1 对所有通信指令进行了优先级分级。

Table 1. The task instruction priority list

表 1. 任务指令优先级表

优先级	任务类型	说明
1	控制	
2	参数设置	
3	保留	优先级 1~5，前台网页操作
4	召测	
5	巡测	
6		
7		
8	自定义	优先级 6~9，自动任务
9		
10	终端升级	耗时操作，优先级最低

确定了各通信任务的优先级，接下来进行任务优先级处理，包括两个方面：不同优先级自动任务之间的调度与同一终端下不同优先级通信任务的处理。处理原则是：自动任务的优先级处理是多个任务先后启动时，先执行优先级高的通信任务，即使是优先级高的任务后启动，也需要优先执行优先级高的任

务，暂停优先级低的任务发送。同一终端多个任务排队执行时，优先执行优先级高的任务，但为了保证通信任务时序性与连续性，当前正在通信的任务不能被高优先级任务中断。

具体实现方案是采用 Kafka 消息队列主题分区功能，合理设定多个指令优先级主题，在发生请求阻塞时，采集前置按照优先级来处理指令，从而满足重要指令的快速下发需求。例如，将费控指令放在优先级为 A 的分区中，召测指令放在优先级为 D 的分区中，当费控下发前排有大量召测请求时，采集前置也会优先处理费控指令。基于 Kafka 消息队列实现优先级划分后，可以实时调整某类请求的优先级别，保证重要命令下发的实时性。

2.2. 采集任务并行执行

当前采集前置中自动任务和 WEB 手动任务全部通过任务调度执行，互相影响效率。

针对此问题，本文将自动任务模块从任务调度拆分出来，专门生成和处理自动任务。任务处理结构如图 1 所示：

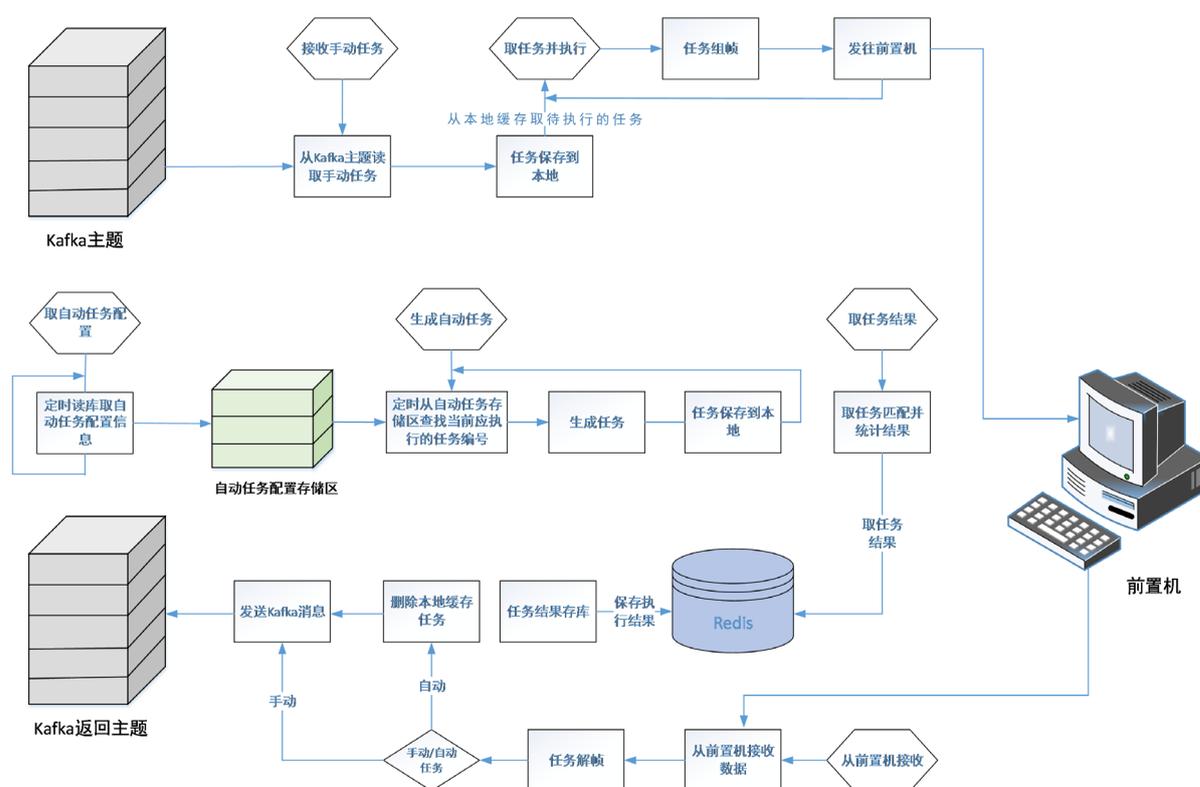


Figure 1. Task processing structure

图 1. 任务处理结构图

数据采集任务执行主要执行模块为自动任务模块、任务调度模块以及前置机模块。自动任务模块主要负责采集任务定时启动与通信任务的生成，该模块的执行效率直接关系到通信任务生成效率和发送效率；任务调度模块是连接应用服务器和前置机的中间系统，担任着任务的接收、生成、参数配置、组帧、传递、解帧、返回发送的角色，主要负责整个采集系统的任务调度工作，由于处理任务的多样性，与应用服务器和前置机交互频繁，会有大量异步操作和线程管理。

为提高任务发送和接收的速度，使用 Kafka 消息中间件，其不仅拥有超高吞吐量、毫秒级的延迟、

极高的可用性和可靠性，而且分布式可以任意扩展。

自动任务集群节点间通信使用 JGroups，使用 TCP 连接，实现多服务器之间的通信，包括维护群组状态、群组通信协议、群组数据可靠性传输等。JGroups 集群中的某个节点会提供在一段时间内维护状态信息和消息可靠性监测的功能。

自动任务采用集群设计但只有一个主系统，主系统比其他系统多了任务分发和任务统计功能，负责任务的启动、运行和监控、结束和取消。主系统的选举和维护通过 Redis 存储，主系统定时更新心跳至 Redis，其他备系统通过读取 Redis 中的心跳信息获悉主系统的信息。

任务调度与前置机的通信采用 socket 连接。为确保同一终端的任务始终发送到同一台前置机，使用终端 ID 跟当前连接的前置机数量取模方式获取发送前置机的连接信息。

2.3. 高并发通信处理

当前采集前置机程序采用 select 方式对并发通信链路进行报文收发管理，前置机进程每次查看是否有数据接收时，需要将所有连接的套接字传送给操作系统，将大量用户态内存拷贝到内核态内存，增加内存开销的同时，由于每次要对所有管理的连接进行遍历，查看是否有数据接收事件发生，造成巨大的资源浪费。

高效并发通信能力是采集系统数据采集与交互能力的根本，针对采集前置机程序采用 select 方式存在的问题，采用 EPOLL 技术来开发管理终端连接处理组件，提升前置机的终端接入及并发通信能力。

Linux 内核的 Epoll 提供了更高效的处理大量并发连接通信的方式，它在 Linux 内核中申请了一个简易的文件系统，将原来 select 调用分成“建立一个 Epoll 对象”、“向 Epoll 对象中添加链接套接字”、“收集发生事件的连接”3 个部分，这样只需要在进程启动时建立 1 个 Epoll 对象，并在需要的时候向它添加或删除连接就可以了。因此，在实际收集事件时，epoll_wait 的效率就会非常高，因为调用 epoll_wait 时并没有向它传递全部的连接，内核也不需要去遍历全部的连接，有效降低内存与 CPU 的使用率。Epoll 适合管理大量连接但同一时刻只有少量连接在应用的场景，因此非常适合数据采集这样的并行通信场景。

Epoll 提供的操作方法 epoll_wait 每次只返回有事件发生(如接收到数据)的连接，并将发生的数量返回，因此 epoll_wait 接收数据比 select 方法高效。

2.4. 规约动态扩展及适配

当前采集前置的规约处理功能与前置服务耦合度过高。一方面使得规约处理模块程序的更新要求采集前置必须重启；另一方面当规约处理模块程序异常时，还可能联动引发其它异常。

针对此问题，将规约处理模块进行组件化拆分，降低规约处理组件间的耦合度，并基于热插拔理念进行组件管理功能的研发，满足单个规约解析包更新不影响整体采集前置服务的要求。

前置机规约库包含在运行所有采集设备、计量设备或感知器的通信规约，为提升通信过程中规约匹配效率，可根据每通信信道所带现场设备实际规约类型进行选配。新增规约时，只需按前置机约定的接口规范开发新的规约插件，将插件放入前置机规约库即可实现使用新规约与现场设备通信。

前置机进行规约适配采用两种规则：一种是根据下行通信任务指定的规约，另一种是根据上行的报文逐个规约插件解析帧结构适配确定规约类型。

前置机记录有过通信交互的终端的规约类型，在主站与现场设备通信时，主站会将档案配置中的规约类型下发给前置机，前置机按指定的规约插件进行报文组装，并记录该终端的规约类型，在终端返回时，就直接调用相同的插件进行解析。

当终端第一次与前置机建立通信链路请求时，前置机根据发送的报文和该通信信道下配置的规约插件逐个解析帧结构，如果帧结构匹配，则认为该协议适用，前置机记录当前终端的通信协议为该规约类

型。除非主站发起通信任务指定的规约类型与识别的规约类型不一样才会修改前置机记录的该终端规约类型，否则将一直使用识别的规约类型与现场设备进行交互通讯。

2.5. 采集数据入库并行处理

当前采集数据一般存储在 Oracle 数据库中，海量高频数据入库会造成数据库瞬时压力过大的问题。数据库瞬时压力过大不仅造成数据入库的不及时，甚至可能导致数据丢失等其他异常，进而影响相关业务的正常开展。

针对高频数据采集引起的高并发入库处理瓶颈，通过重构系统存储体系来解决处理。引入分布式关系库存储框架，搭建大规模并行处理能力的分布式关系库(作为高频数据，如 15 分钟一个点的曲线采集的存储)，如图 2 所示。

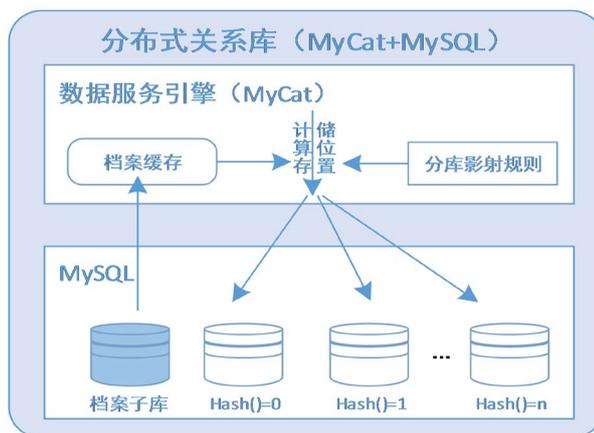


Figure 2. Distributed relational database
图 2. 分布式关系库

分布式关系库使用 MyCat 中间件 + MySQL 数据库构建 MPP (Massively Parallel Processor) 数据库，采用分库分表存储方式，实现数据高效智能的水平拆分与分布式存储。例如，分布式关系库分布在 8 台 MySQL 主机的 16 个分片上，通过 mycat rule.xml 设定分片规则，通过 schema.xml 集成各个 DataNode (分布式 MySQL 数据库) 来构建逻辑库和逻辑表从而供应用使用，如以正向有功示值为例，该表采用 ID mod 16 的分片规则，将 16 个 MySQL 数据库上的实际存在的表格构建成一个虚拟的逻辑表，应用直接使用该逻辑表进行入库，入库时套用分片规则，根据 ID mod 16 的规则将数据均匀的存储到 16 个数据节点上。利用分布式关系库通过分库分表以及扩展节点增加并发处理能力的特点，缓解短时间周期内大量数据入库的压力，满足了大规模高频采集数据入库时高并发处理能力的需求。

3. 应用验证

基于以上关键技术，对采集前置进行优化提升，在国网某省级电力公司用电信息采集系统主站提升改造项目中，开展了工程应用。采集前置服务部署运行在由 50 多台服务器组成的多节点、集群环境上。服务器基本配置为 CPU：双路 12 核，2.1 GHz；内存：128 G。

该省电力公司采集系统接入 38 万台集中器及 40 万台公专变采集终端，覆盖 2700 多万低压用户。此次验证选取了全量低压用户日冻结示值数据采集、高压用户 96 点冻结曲线数据采集 2 个业务应用场景，进行了 2 周的观察，计算平均性能。采集前置优化改造前后在终端接入及通信、数据采集及入库方面平均性能对比如表 2 所示。

Table 2. The performance of acquisition front-end system before and after transformation
表 2. 采集前置优化改造前后性能对比

指标项	优化前	优化后
单台前置服务器支持承接连接数	2 万	10 万
单台前置服务器并发通信能力	1 万	3 万
信道、设备正常情况下，与终端交互单条报文响应时间	≥3 秒	≤1 秒
终端上行报文解析平均速率	≤6 万条/秒	≥10 万条/秒
全量低压用户日冻结示值数据采集	62.3 分钟/次	11.9 分钟/次
高压用户 96 点曲线数据采集	26.3 分钟/次	9.8 分钟/次
数据入库能力	≤1.5 万条/秒	≥9 万条/秒

通过结果对比，可以看到优化改造后的采集前置系统在终端接入及通信、数据采集及入库方面的性能有了明显的提升。

4. 结语

本文分析了目前采集前置存在的性能问题，提出了一套解决采集前置性能瓶颈的整体解决方案，涵盖通信任务分级管理、采集任务并行执行、高并发通信处理、规约动态扩展及适配、采集数据入库并行处理等多个方面。在国网某省级电力公司采集系统主站的提升改造工程应用验证中表明，提高了采集前置服务的接入能力和并发处理能力，满足了高效数据采集的性能要求，对国网各省采集系统主站的升级改造具有一定的借鉴意义。在后续的工作中将结合物联管理平台架构做进一步研究完善。

参考文献

- [1] 何恒靖, 赵伟, 黄松岭, 等. 云计算在电力用户用电信息采集系统中的应用研究[J]. 电测与仪表, 2016, 53(1): 1-7.
- [2] 张治, 高曦莹, 李广翱, 等. 基于云平台并行分布式计算对电力采集业务性能提升的研究[J]. 计算机科学与探索, 2017, 11(增刊): 397-404.
- [3] 刘振亚. 全球能源互联网[M]. 北京: 中国电力出版社, 2015.
- [4] 华煌圣, 甘霖, 伍竹林, 等. 基于物联网技术的能源互联网数据支撑平台[J]. 供用电, 2018, 35(1): 40-45.
- [5] 王赞, 陈光, 董晓, 等. 基于工业互联网的智慧能源服务系统架构研究[J]. 电力系统保护与控制, 2020, 48(3): 77-83.
- [6] 武昕, 焦点, 高宇辰. 基于非侵入式用电数据分解的自适应特征库构建与负荷辨识[J]. 电力系统自动化, 2020, 44(4): 101-109.
- [7] 宋永华, 包铭磊, 丁一, 等. 新电改下我国电力现货市场建设关键要点综述及相关建议[J]. 中国电机工程学报, 2020, 40(10): 3172-3186.
- [8] 陈艳, 宋英华. 新型配电物联网后台系统架构设计与关键技术研究[J]. 供用电, 2020, 37(2): 41-46+53.
- [9] 宋祺鹏, 石岩, 韩子龙, 等. 泛在电力物联网下的配电网典型配置模式研究[J]. 供用电, 2020, 37(1): 15-20.
- [10] 王林童, 赵腾, 张焰, 等. 配用电大数据多源集成及存储优化方法[J]. 高电压技术, 2018, 44(4): 1131-1139.
- [11] 熊霞, 陶晓峰, 高鲁鑫, 等. 综合能源一体化采集系统的多任务自适应实时调度方法[J]. 电测与仪表, 2019, 56(20): 108-114.
- [12] 张杰, 刘凯, 周立军. 采用 Redis 高并发应用系统设计与实现方法[J]. 计算机与数字工程, 2020, 48(5): 1222-1226.
- [13] 邬林. 基于消息队列的用电信息采集技术的研究与实现[D]: [硕士学位论文]. 北京: 华北电力大学, 2019.
- [14] 赵玉文. 面向物联网的自适应异构数据采集技术的研究与实现[D]: [硕士学位论文]. 北京: 北京邮电大学, 2014.