

DOI:10.3969/j.issn.1003-5060.2024.08.004

基于 OPC UA 的分布式数据采集处理系统架构研究

扈静, 柏晨, 张玺, 葛茂根

(合肥工业大学 机械工程学院, 安徽 合肥 230009)

摘要:随着物联网与互联网融合的不断深化,感知层与应用层之间的互联互通要求不断提高。针对现有数据采集系统存在的规范性、扩展性和适应性不足的问题,文章提出一套基于对象连接与嵌入过程控制统一架构(object linking and embedding for process control unified architecture, OPC UA)协议的分布式数据采集处理系统架构。对数据采集处理系统各个模块进行功能解耦和架构重组,分别介绍该系统硬件和软件架构;针对发布订阅模式下负载的特异性,提出一种改进的适用于 OPC UA 分布式订阅的负载均衡算法;最后在某车企实例验证该系统架构。结果表明,基于 OPC UA 的分布式数据采集处理系统数据采集处理效果良好,证明了该系统架构的可行性及有效性。

关键词:系统架构;分布式;基于对象连接与嵌入过程控制统一架构(OPC UA)协议;负载均衡

中图分类号:TP274.2 **文献标志码:**A **文章编号:**1003-5060(2024)08-1028-07

Research on the architecture of distributed data acquisition and processing system based on OPC UA

HU Jing, BAI Chen, ZHANG Xi, GE Maogen

(School of Mechanical Engineering, Hefei University of Technology, Hefei 230009, China)

Abstract: With the continuous deepening of the integration of the Internet of Things and the Internet, the interconnection between the perception layer and the application layer is increasingly required. In order to address the deficiencies of the existing data collection system in terms of specifications, extensibility and adaptability, this paper puts forward a distributed data acquisition and processing system architecture based on the object linking and embedding for process control unified architecture(OPC UA) protocol. Firstly, each module of the data acquisition and processing system is decoupled and re-organized, and the hardware and software architectures of this system are introduced. Then, an improved load balancing algorithm for OPC UA distributed subscription is presented according to the load specificity in publish/subscribe mode. Finally, this architecture is verified in an example of a vehicle enterprise. The results show that the distributed data acquisition and processing system based on OPC UA has good data acquisition and processing effect, which proves the feasibility and validity of the system architecture.

Key words: system architecture; distributed; object linking and embedding for process control unified architecture(OPC UA) protocol; load balancing

信息技术高速发展不断推进物联网各类关键技术的迭代和升级,在“智能制造 2025”不断临近的背景下,实现工业系统信息化、智能化的一个重

要课题是形成完备的工业信息采集和控制系统。实现底层自动化控制设备的数据同上层制造执行系统(manufacturing execution system, MES)和

收稿日期:2022-12-14

基金项目:安徽省科技攻关计划资助项目(JZ2016AKKG0837)

作者简介:扈静(1976—),女,安徽六安人,博士,合肥工业大学副教授,硕士生导师。

企业资源计划(enterprise resource planning, ERP)深度交互,构建全流程透明化协同生产,是构建智能制造工厂的重要步骤^[1]。目前离散型制造业数据采集系统主要存在以下问题。

1) 底层自动化工业设备品牌、种类繁多,通信协议和规范不统一,在车间设计规划阶段未预留规范接口,设备与系统之间通信扩展性较差。

2) 现有的数据通信实现形式较为冗杂,难以集成的数据往往通过增加中间数据或接口的形式实现集成,导致系统数据冗杂利用率低,同时造成系统处理速度下降,系统性能降低^[2]。

3) 部分企业虽然采用了基于对象连接与嵌入过程控制统一架构(object linking and embedding for process control unified architecture, OPC UA)协议的数据采集方案^[3],但大多采用单体系统形式实现,系统性能受限于硬件设备、吞吐量,存在一定瓶颈。

4) 现有的工厂数据采集系统的部署和实现与当前互联网发展融合度较低,成果转化难度较大。

针对以上问题,本文研究基于 OPC UA 的分布式数据采集处理系统架构,基于 OPC UA 特点将数据采集处理系统分为配置中心、采集处理中心、可视化监听 3 个模块。

1) 配置中心模块。一方面综合配置需要采集和控制的数据,实现系统热更新;另一方面实现客户端权限的管理,提高数据安全性;除此之外,所有外部控制请求将由配置中心接受并进行权限

与安全性验证,符合条件的设备操作通过配置中心调用采集处理中心后端接口,实现对设备的控制。

2) 采集处理中心模块。采集处理中心硬件部分通过可编程逻辑控制器(programmable logic controller, PLC)集成底层设备数据,软件部分负责与硬件层通过 OPC UA 协议进行直接数据交换^[4],通过配置模块实现对监听对象的监听和实时控制,采用改进的适用于 OPC UA 分布式订阅的负载均衡算法,实现采集模块性能的最大化;同时为业务处理预留接口模块,采用集群部署加负载均衡的方式增加系统吞吐量,提高系统的响应速度,预留各类接口并暴露给配置中心层,以便系统功能扩充。

3) 可视化监听模块。基于前 2 个模块的数据综合进行可视化展示,实现数据采集交互全流程各个节点状态的监控,提高生产流程透明度,通过对历史数据进行检索计算形成各类生产数据报表,提高全流程生产数据利用率。

本文设计的系统架构软件方面基于 Kubernetes 和 Docker 技术实现,并采用分布式集群部署的形式,降低系统的耦合性,提高系统的性能。

1 系统整体架构设计

1.1 采集模块硬件架构设计

为提高系统鲁棒性以及相应速度,本文基于 OPC UA 的分布式数据采集处理系统采集模块硬件架构如图 1 所示。

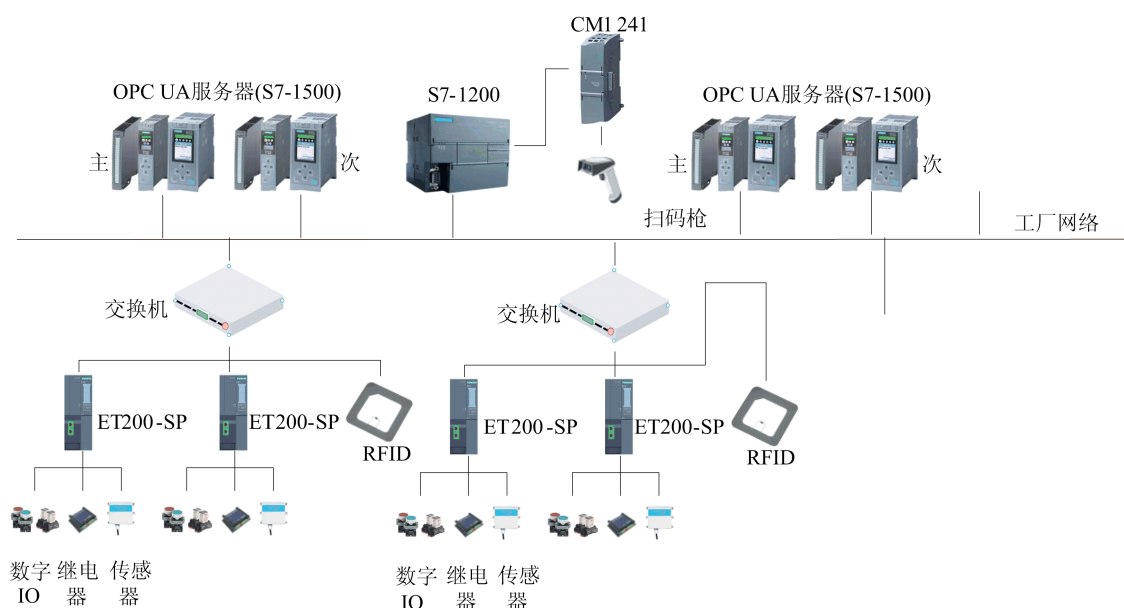


图 1 基于 OPC UA 的分布式数据采集处理系统硬件架构

工业现场设备众多、布局较为复杂,为了高效集成和管控现场众多设备,本文采用图 1 所示的拓扑结构。以支持 OPC UA 的 PLC 作为 OPC UA 服务器(图 1 中 S7-1500),服务器 PLC 一方面通过 Profinet 直接集成现场设备(如 IO、继电器、传感器、光电对射开关、无线射频识别等)信号^[5];另一方面先通过单元 PLC(图 1 中 S7-1200)采集部分模块信号,再通过 PN/PN 耦合器(PN/PN Coupler)实现 PLC 间通信,通过以太网将设备数据分模块采集到分布式 OPC UA 服务器中。每组 OPC UA 服务器采用主次模式和交叉采集模式部署,每个为外部提供 OPC UA 服务的 PLC 交叉采集所属模块下所有数据,均向外提供心跳节点。部分单元 PLC 不需要提供 OPC UA 服务,故可采用型号较低的 PLC,缩减成本。在系统完善的前提下,主服务器 PLC 向外提供订阅服务,所有数据跳变信息由 PLC 向外发布,次 PLC 向外提供读写服务,软件端逻辑处理下发的各项指令,由次服务器 PLC 提供。通过心跳信号实现服务器故障情况下的即时切换,增加系统鲁棒性,减少因服务器故障造成的大批设备失控掉

线而造成的损失,为检修提供时间。

OPC UA 服务器 PLC 群组通过以太网连接,采用分布式集群部署形式向软件端提供 IP 和端口,工厂内部众多设备单元分布式部署,单一单元功能受损不影响其他单元数据的交互,并且能够缩短定位检修时间。模块内部采用集群部署方案,至少采用主、次 2 个 OPC UA 服务器,主服务器负责信息交互量较大的订阅服务,次服务器提供读写控制服务,提高系统吞吐量的同时,能够使 2 个集群互为备用。软件模块只需要接入工厂网络即可通过 IP 加端口与 OPC UA 服务器进行连接通信,与设备层进行数据交互。

1.2 分布式软件系统架构设计

软件系统后端采用 Java、前端采用 Vue 开发,并基于 Kubernetes 和 Docker 技术实现分布式集群部署,关系型数据库采用 MySQL,非关系型数据库采用 Redis。

软件端分为配置中心、采集处理中心、可视化监听 3 个模块,其中配置中心和可视化监听模块分前、后端模块,采集处理中心仅后端服务。软件系统功能架构如图 2 所示。



图 2 软件系统功能架构

配置中心负责管理和配置采集处理模块 OPC UA 客户端权限和功能并实现数据持久化。通过配置模块统一管理所有 OPC UA 服务器的权限,同时为具备权限的客户端配置订阅和控制的数据节点名称,并通过接口实现采集模块热更新;所有外部控制请求将直接请求到配置中心后端,由配置中心统一认证权限以及操作安全性,确保安全的操作指令才能发送到硬件端。配置中心接收外部指令并判断安全后,调用采集处理中心

后端接口进而对设备端进行指令下发。除采集处理中心服务外,其他任何服务或接口不可直接对硬件进行操作。

采集处理中心为核心业务模块,负责与硬件层进行数据交互。该服务模块通过 OPC UA 协议监听 OPC UA 服务器中信号点的变化,并通过改进的负载均衡算法,将跳变信号合理分配到不同集群中进行处理,最终完成功能的闭环与数据持久化。由于 OPC UA 订阅的处理模式不同于

普通服务间的 API 接口调用,无法复用微服务框架中负载均衡的实现方案(如通过 Nginx 实现接口调用的负载均衡),故本文采用一种改进的适用于 OPC UA 分布式订阅的负载均衡算法,对由底向上的跳变处理请求进行合理分配,使处理系统达到最优化资源使用、最大化吞吐率、最小化响应时间,同时避免过载。该模块集成模块化设备控制接口,通过 OPC UA 的短时连接写入功能实现信号接收与控制指令下发的解耦,同时保证系统可扩展性。

可视化监听模块服务提供人机交互界面,通过配置实现底层数据实时可视化展示,实现方式

为通过 WebSocket 与采集处理模块后端之间建立长时对话通道,实现数据采集交互全流程监控。在满足发送条件时,采集处理模块后端通过 WebSocket 直接将处理好的数据发送到可视化监听前端服务模块,前端通过 Vue 的 v-bind 指令实现前端组件的实时动态变化,避免前端为刷新数据而进行反复长时轮询操作。另外,该模块负责输出工厂全流程数据报表,集成数据查询功能与数据整合计算功能,实现多元数据展示,提高工厂全流程数据使用效率。

本文系统各个服务之间分工明确,分布式软件系统架构如图 3 所示。

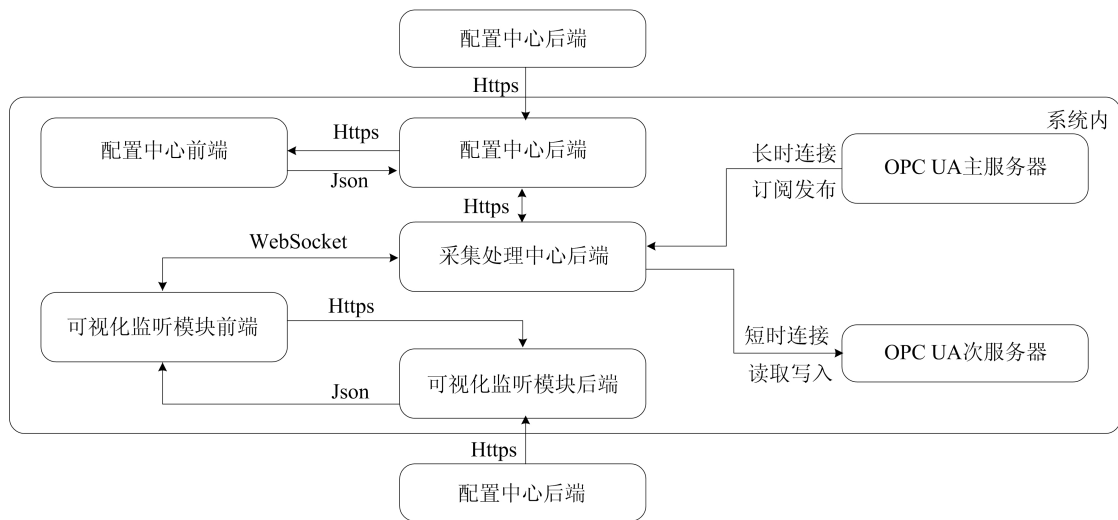


图 3 分布式软件系统架构

配置中心涉及工厂数据安全与权限控制,目标用户为具备一定电气知识的管理人员,其关注于硬件采集层之上的软件采集层的工作情况,对整个工厂的不同数据权限进行划分,并且可以通过接口直接控制采集处理层开始或终止对某些数据的采集处理流程。其中:前端服务提供可视化操作页面;后端服务主要实现权限相关数据的数据持久化以及各类相关接口的调用。配置中心作为统一对外门户,还处理其他系统对硬件的操作请求,认证后再将请求转发到采集处理中心后端。

采集处理中心作为整个模块的核心,其核心职责是直接和硬件层众多分布式 OPC UA 服务器建立连接进行数据交互,并完成工厂硬件数据记录相关的数据持久化。作为直接进行数据交互的模块,其受限于可视化监听模块的权限和指令控制,并将所采集到的数据发送到可视化监听模块前端完成实时展示。

可视化监听模块前端数据来源有 2 种形式:

一种是通过 WebSocket 直接从采集处理中心后端获取实时数据,无需经过数据库查询;另一种是通过请求可视化监听模块后端检索并计算出所需的数据或报表。此外可视化监听模块后端也作为其他模块数据交互的中转站,所有数据库数据记录查询的相关请求均由可视化监听模块后端提供支持。

2 关键使能技术

2.1 基于分布式集群部署的系统架构

通过 Kubernetes 和 Docker 技术实现的分布式集群部署与传统制造业 MES 系统的单机单服务器部署方式存在巨大差异,它不仅具有系统健壮性强、可扩展性高、更易于管理和升级等优点,还能够将现成的互联网微服务发展的成果进行快速转化,是未来企业智能制造系统发展的趋势所在。

1) 系统采用分布式集群部署方式进行设计,将复杂的业务拆分为多个较小的业务,每个服务

专注于本身的功能,能够更好地实现功能复用。每个服务独立部署、独立扩容,将单个服务更新和升级对其他服务造成的影响降到最低。将服务拆分后,伴随着功能细化的同时,也将错误和故障隔离开,一个服务中的问题不会影响到整个应用程序,而将被限制在某一个服务之中,其他服务与之没有交互的功能将正常运行。这不仅减少了错误所带来的损失,也为寻找问题的根源提供了方向。不同服务之间独立开发、独立部署,进行任务划分时也将更加合理,团队开发和维护效率得到提升。同时 OPC UA 服务器也采用分布式部署,不同模块数据汇总到不同的 OPC UA 服务器,单服务器的故障或升级不影响其他模块,同样具备上述分布式软件系统所具有的优势。

2) 软件系统采用 Kubernetes 和 Docker 技术实现分布式集群部署^[6]。Docker 是一个开源的应用容器引擎,能极大程度地消除开发与运维环境的差异,便于进行版本管理和持续交付,使开发者既能将其应用和依赖包打包到一个可移

植的容器中并发布到任何流行的机器上,又可以实现虚拟化。Kubernetes 是一个开源的容器集群管理系统,可以实现容器集群的自动化部署、自动扩缩容以及维护等功能。具体使用步骤如下:将 Web 应用打包为 JAR 并使用内部的 JAR 文件创建 Docker 映像,创建好的 Kubernetes 部署后,Kubernetes 使用 Docker 映像启动 Pod,最后创建一个 Kubernetes 服务作为所有正在运行的 Pod 的入口点,完成 Kubernetes 集群对应用的托管。

2.2 基于分布式订阅的负载均衡算法实现

应用程序编程接口(application programming interface, API)调用所需要实现的负载均衡主要是对接口调用请求根据服务器资源情况进行分配,以提高服务的性能和可靠性^[7]。不同于传统 API 访问的由服务外向服务请求处理,本文系统 OPC UA 采用订阅发布模式,采集处理中心服务采用集群部署并均与所有 OPC UA 服务器组建立连接,如图 4 所示。

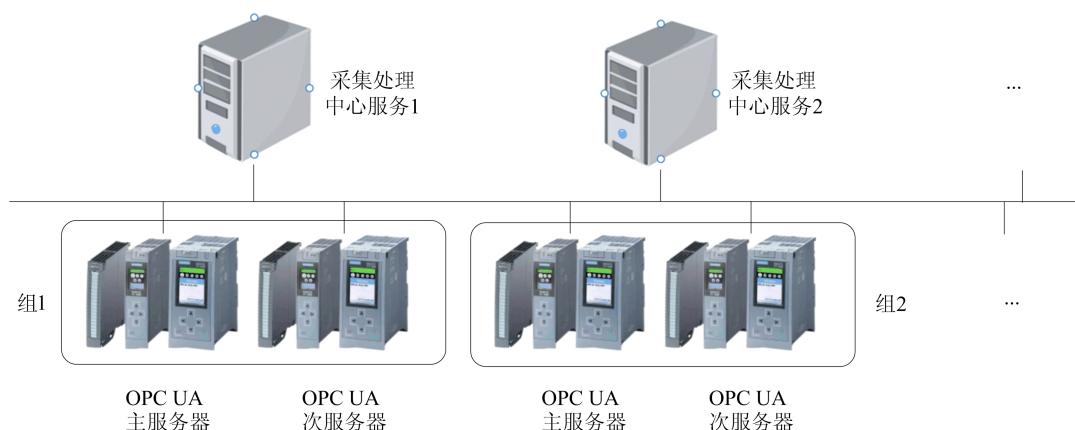


图 4 集群部署架构图

采集处理中心 n 个集群通过 OPC UA 协议,使用 IP 加端口与所有分布式 OPC UA 服务器群(即支持 OPC UA 的 PLC,下文简称 PLC)建立连接,订阅服务器中所需数据节点,设备端数据变化时,对应 PLC 节点内数值发生变化,订阅对应节点的所有客户端则会收到本次跳变信息。根据 OPC UA 订阅发布模式^[8],OPC UA 服务端以一定周期扫描被订阅节点,发生数值变化则发送节点名称、节点跳变时间点和节点内数据值等给到 OPC UA 客户端,因此建立连接和接收跳变信息对采集处理中心后端性能消耗较小,对应节点跳变后的逻辑处理对于性能消耗较大。本文系统采用交叉采集后进行负载均衡处理的形式,将采集

处理中心 n 个副本订阅所有 PLC 内数据节点,实现方式如下所述。

1) n 个采集处理中心后端全量订阅所有 PLC 内数据节点,任意节点跳变时,所有采集处理中心后端均可收到跳变信息(包括节点名称、节点跳变时间点和节点内数据值)。

2) 为各个服务定义一个负载压力值的概念,反映当前任务对服务的性能消耗压力情况,公式如下:

$$P_i = Q_i + N_i X \quad (1)$$

其中: P_i 为当前服务的负载压力值, P_i 值越大,代表当前服务负载越大,接收新任务能力越弱; Q_i 为空运行情况下每个服务负载压力初始值,即当

所有服务均未开始处理节点跳变数据时,性能越好的硬件对应的服务空压力越小,接收数据处理请求的优先级越高; N_i 取值为各服务对应性能指标因素,由服务对应的硬件性能决定,性能越好, N_i 越小, N_i 取值范围为(0,1]; X 为服务当前正在处理中的请求个数。

3) 通过对各服务对应服务器的性能对比(本文是对各个 Docker 容器性能进行对比),初步给定各服务器 Q_i 和 N_i 值,为保证 Q_i 仅代表 X 为 0 时各服务性能差异,规定 Q_i 数量级为 $\min(N_i)$ 数量级的千分之一。例如当前有 2 个性能相同和 1 个性能大致为前 2 个性能 2 倍的服务(3 个服务分别称为 S_1 、 S_2 、 S_3),可取 $N_1 = N_2 = 1, N_3 = 0.5$,取 $Q_1 = Q_2 = 0.000 2, Q_3 = 0.000 1$ 。当所有服务均无处理任务时, S_3 负载压力值最小,优先处理任务;当 S_3 已接收一个任务正在处理时,若当前 S_1 、 S_2 负载压力值均小于 S_3 ,则由 S_1 或 S_2 进行处理,依次类推。

4) 所有采集处理中心后端共用一个 Redis 非关系型数据库,将所有服务器对应的服务器名与当前服务的负载压力值以键值对的形式存放在 Redis 中。又因任意次 OPC UA 订阅跳变,均可用(节点名称,节点跳变时间点)唯一确定,亦可用 Redis 实现分布式锁。故当所有后端接收到当前次跳变时,首先根据节点名称加节点跳变时间点判断当前任务是否已被处理,若不存在已被其他服务处理的标识,则判断自身负载压力值是否最小,若两点均满足,则将本次跳变信息传递到负责数据处理部分的函数中,将节点名称和节点跳变时间点在 Redis 中联合标记为已处理,并在 Redis 中更新自身当前负载压力值。处理完成后,删除本条标记数据并更新负载压力值即可,不满足处理要求的服务则直接放弃处理本次请求。算法处理流程如图 5 所示。

5) 基于预先设定的参数,提供 2 种优化渠道:① 通过人为更改参数 N_i 和 Q_i 值调优;② 通过服务运行若干时间后,统计服务任务平均处理时长,取反比后进行归一化,即可获得各个服务器对应的 N_i 值,通过对 N_i 值进行升序排序可以得到 Q_i 值计算公式,即

$$Q_i = \frac{mN_{\min}}{1000} \quad (2)$$

其中: N_{\min} 为各服务器任务平均处理时长取反比后归一化所得值中最小值; m 为各服务器 N_i 值升序排序后当前服务器对应的次序。

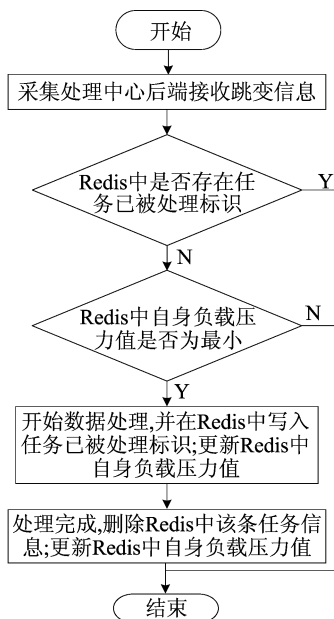


图 5 算法处理流程

3 系统架构验证

本文 PLC 程序采用博图编写,软件程序后端采用 Java、前端采用 Vue 开发。以 SIMENS S7-1500 PLC 作为 OPC UA 服务器,PLC 设置如图 6 所示。

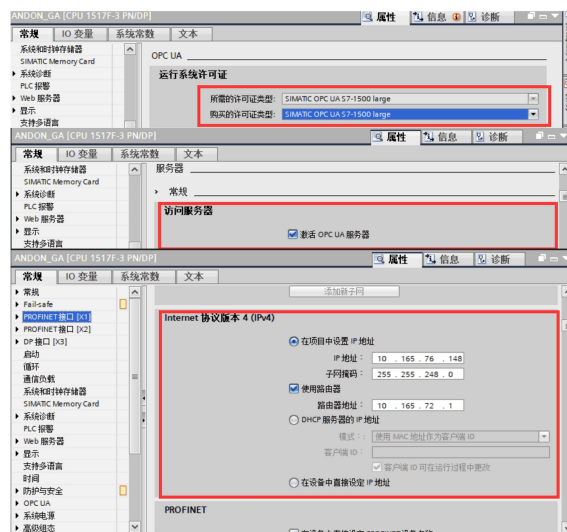


图 6 OPC UA 服务器 PLC 设置

PLC 集群接入工厂网络后与设备建立连接,采集处理中心根据配置中心权限,通过 IP 加端口与 PLC 群建立连接实现数据采集,并将数据发送至可视化监听模块前端,展示效果如图 7 所示。

另外,本文系统还提供了外部硬件操作服务、数据报表服务、外部数据调用服务等功能,保证兼容性和可扩展性。该系统已在某大型车企运行半

年,数据采集处理效率良好,证明了系统的可行性以及有效性。



图 7 可视化监听模块效果展示图

4 结 论

本文提出的基于 OPC UA 的分布式数据采集处理系统软件端架构与互联网微服务架构较类似,易于将各类互联网微服务架构中各类优势技术与工具融合。考虑到数据采集处理系统不是完整意义上的全工厂智能制造系统,本文系统未引入微服务架构中的各类技术,如后续扩充系统功能,可将各类互联网成果转换运用。例如采用统一的网关服务管理所有前端请求或外部请求,作为暴露服务的唯一出口,所有 API 请求优先进入网关服务,通过微服务网关实现用户验证、访问权限控制、过滤拦截、限流、熔断等功能^[9];增加各类辅助性基础服务(文件、消息、字典服务等),通过更多精细化服务使整个系统更加细致高效。在微服务构建的系统内部,采用注册中心和配置中心(如 Nacos)实现服务注册、服务发现和 API 负载均衡等。

现代制造业智能化、信息化需要与互联网的发展相互借鉴,在智能制造设计的全过程中,立足本身特点,合理运用互联网各项先进技术,将为智能制造系统发展注入新的动力。

[参 考 文 献]

- [1] 周小雯,杨磊,袁明锋. 基于 ERP 和 MES 平台协同的智慧云排产方法研究[J]. 中国仪器仪表, 2021(8): 33-38.
- [2] 葛宁. 基于 OPC UA 的智能车间数据采集与监控系统[D]. 大连: 大连理工大学, 2021.
- [3] 禹鑫焱,殷慧武,施甜峰,等. 基于 OPC UA 的工业设备数据采集系统[J]. 计算机科学, 2020, 47(2): 609-614.
- [4] HAN Y J, HU Y M, WANG Y Q, et al. Research and application of information model of a lithium Ion battery intelligent manufacturing workshop based on OPC UA[J]. Batteries, 2020, 6(4): 52.
- [5] 李玉敏. PROFINET 技术在工业数字化转型中的应用探讨[J]. 电气时代, 2022(3): 6-10.
- [6] 陈志达. 基于 Kubernetes 和 Docker 技术的航油容器云平台架构设计[J]. 电子技术与软件工程, 2021(24): 148-149.
- [7] 经成,谢军. 基于微服务架构的负载均衡优化算法及实现[J]. 南昌大学学报(理科版), 2022, 46(4): 399-404.
- [8] PU C G, DING X W, WANG P, et al. Semantic interconnection scheme for industrial wireless sensor networks and industrial internet with OPC UA Pub/Sub[J]. Sensors, 2022, 22(20): 7762-7762.
- [9] 廖俊杰,陶智勇. 微服务 API 网关的设计及应用[J]. 自动化技术与应用, 2019, 38(8): 85-88.

(责任编辑 胡亚敏)

(上接第 1027 页)

- [19] 秦萍萍,朱德成,贾晓晨,等. 磁絮凝工艺处理初期雨水的研究[J]. 环境科学与技术, 2019, 42(12): 144-148.
- [20] LV M, ZHANG Z, ZENG J, et al. Roles of magnetic particles in magnetic seeding coagulation-flocculation process for surface water treatment[J]. Separation and Purification Technology, 2019, 212: 337-343.
- [21] LIU C, WANG X, QIN L, et al. Magnetic coagulation and flocculation of a kaolin suspension using Fe₃O₄ coated with SiO₂[J]. Journal of Environmental Chemical Engineering, 2021, 9(5): 105980.
- [22] 魏超. 磁絮凝-吸附组合工艺去除污染河水氮磷效果研究[D]. 无锡: 江南大学, 2021.
- [23] MA J, XIA W, FU X, et al. Magnetic flocculation of algae-laden raw water and removal of extracellular organic mat-

ter by using composite flocculant of Fe₃O₄/cationic polyacrylamide [J]. Journal of Cleaner Production, 2020, 248: 119276.

- [24] 郑利兵,吴振军,张鹤清,等. 介质加载混凝过程中磷和溶解性有机物的去除特征及混凝机理[J]. 环境工程学报, 2020, 14(12): 3352-3362.
- [25] LV M, LIU T, CHEN F, et al. Interactions between magnetic particles and polyaluminum chloride on the coagulation behavior in humic acid-kaolin synthetic water treatment[J]. Environmental Research, 2021, 197: 111093.
- [26] 罗强,张君,李晔,等. 复合混凝剂去除湖泊水中磷的研究[J]. 武汉理工大学学报, 2020, 42(7): 50-54.

(责任编辑 吴亮)