TongETL技术指标

**V24.2.6**

|  |  |  |  |
| --- | --- | --- | --- |
| **编写部门** | 技术研究部 | **编制人员** |  |
| **审核人员** |  | **批准人员** |  |
| **发布日期** | 2024.1 | **发布范围** | 公司内部 |
| **版本** | V1.0 | **文件密级** | 公开 |

# 功能指标

|  |  |
| --- | --- |
| **类别** | **技术指标** |
| 总体要求 | 根据国家有关部门相关政策、本系统建设的技术需求以及信息安全性要求，要求投标产品应具有国内自主知识产权，并提供产品软件著作权证书。 |
| 产品架构成熟、功能稳定，在全国大型的多级部署的数据交换项目中具有成功应用案例，提供5个以上部省级数据交换平台项目案例。 |
| 适配国产化环境，详见附录 |
| 数据交换功能 | 基于标准的JDBC接口，实现对各种主流数据库系统的支持。详见附录 |
| 提供丰富的数据文件抽取和加载组件，支持包括普通文本、CSV、XML、Excel、Json等多种格式的文件。 |
| 支持HTTP、JMS、FTP、Web Services等协议和其他应用系统进行交互。 |
| 提供与大数据生态圈交互适配能力，如：Hive，HDFS，HBase，Kafka，MPP(Postgres-XC/XL)、Redis、MongoDB等。 |
| 提供丰富的数据处理组件，包括值替换、计算、排序、合并、拆分、内容过滤、去重、路由、行列转换、关联查询等处理组件，以及数据校验、字段选择、记录生成、字段值常量/随机数生成工具组件。 |
| 支持全量交换，增量交换，增量交换包括（触发器/时间戳/标志位），支持双向数据同步。 |
| 提供基于增量数据期的重跑、补采场景的支持，支持串行、并行补采配置，提供重跑批次执行参数调整。 |
| 支持基于元数据进行流程输入输出配置，元数据支持关系数据库、Mongo、HBase、HDFS、结构化、半结构化文件等 |
| 任务流程提供多个调度机制和异常恢复机制，在异常后，支持自动和手动恢复。异常恢复可以保证恢复的流程从异常点开始重新同步，保证数据的最终完整性和一致性。 |
| 提供图形化的数据交换配置界面，使用人员无需编码，既可以灵活、方便地定制出各种数据集成流程。 |
| 支持集群部署，流程可在集群内自动分配，服务器伸缩后可平滑切换 |
| 支持故障恢复，集群的服务器出现故障后，系统会自动尝试拉起流程 |
| 支持任务流程和转换流程不依赖资源库运行 |
| 提供批量配置，可通过模板快速的配置和部署数据交换流程，实现批量配置流程和部署流程的功能。 |
| 场景服务 | 提供典型交换场景服务能力，内置数据交换和接口数据采集类场景，如典型的同构库数据备份、异构库同构表的数据全量和增量同步、异构系统文件传输等 |
| 提供流程模版定义能力，可抽象一类场景功能做为模版，方便批量使用 |
| 支持同构API接口调用实现流程动态运行 |
| 流程服务 | 通过服务的形式对外部系统提供流程处理能力。用户可将转换流程部署为流程服务，通过外部API调用实现任务调度，系统自动为每一个服务调用生成一个服务实例 |
| 将流程中可变环节进行参数化，通过变量即可实现业务逻辑，实现用一个流程满足多个同类业务需求的场景,而无需重复定义多个转换流程 |
| 流程调度 | 支持周期性调度、消息调度、分片调度、一次性等流程调度方式。 |
| 提供多样化的调度策略，包括 Cron定时调度、固定间隔调度、手动调度（立即运行）、阻塞策略（串行、并行）、重试策略、调度排序机制、超时策略、数据偏移配置、依赖检查调度等 |
| 监控管理 | 提供统一监控管理工作台，便于全局掌握系统运行情况 |
| 支持对运行的集群、服务器、数据源等基础设施运行情况的监控 |
| 支持任务流程和转换流程的远程状态监控、启动、停止、删除和修改调度周期、日志查看、修改服务器配置【界面和API接口】 |
| 支持流程运行的错误日志单独输出至指定文件，与普通运行日志分离 |
| 支持将数据库连接池共享部署到服务器上，以便实时修改数据库连接池配置 |
| 支持全流程跟踪及展示，可按步骤或组件进行细粒度的跟踪，支持实时运行情况监控和快照统计 |
| 支持ETL流程排期监控，能够直观显示未来一段时间内的任务排期情况，可细化到具体每天某个时段的任务排期计划 |
| 支持执行情况监控，能够查看项目下各个流程的执行状况统计。 |
| 支持基于自定义对账规则的转换执行情况对账，便于从业务上判断任务执行结果是否符合预期 |
| 提供监控管理接口，方便用户快速将监控管理功能集成到自己的平台中 |
| 数据安全 | 基于项目进行资源隔离，包括数据源的使用范围和使用权限，软件的功能权限等 |
| 能够根据租户对数据源进行逻辑划分，以便控制不同业务对数据源的局部访问权限；能够细化到库、表的权限控制 |
| 提供基于角色的权限管理，支持安全员、审计员、管理员三权分立，提供功能级、数据级授权，可以按照使用平台需求，自定义管理员、业务人员、运维人员等不同角色和自定义用户的权限 |
| 多人协作与权限控制，支持流程共享、流程检入检出、版本管理和回滚等 |
| 支持数据传输过程中的数据加密、脱敏等安全措施，支持至少5种脱敏规则 |

# 附录

## 数据库适配清单

|  |  |  |  |
| --- | --- | --- | --- |
| 数据库 | 版本 | 标签 | 备注 |
| MySQL | 5.1，5.7 ，8 | 关系型 |  |
| Oracle | 11.3 | 关系型 |  |
| SQL Server | 2019 | 关系型 |  |
| DB2 | 11.5 | 关系型 |  |
| PostgresSQL | 11 | 关系型 |  |
| GreenPlum |  | 关系型 |  |
| 达梦数据库(DM) | 8 | 关系型 | 国产 |
| 南大通用数据(GBase） | 8s | 关系型 | 国产 |
| 华为高斯数据库 | 6.5 | 关系型 | 国产 |
| 人大金仓数据库 (Kingbase) | V8 | 关系型 | 国产 |
| 神通数据库 | 7.x | 关系型 | 国产 |
| 瀚高数据库 | V5 | 关系型 | 国产 |
| 优炫 | 3.5 | 关系型 | 国产 |
| 航天昆仑 | 3.2 | 关系型 | 国产 |
| 虚谷 | 11.0.0 | 关系型 | 国产 |
| Sybase | ase16.0 | 关系型 |  |
| HBase | 2.1.10 | 列式 |  |
| MongoDB | 3.2.2 | 文档型 |  |
| ES | 5.x、8.x | 文档型 |  |
| MinIO |  | 对象存储 |  |