

Table of Contents

Introduction	1.1
0.概述	1.2
1.快速开始	1.3
1.0 常规部署	1.3.1
1.1 docker镜像部署	1.3.2
1.2 常规部署	1.3.3
2.适用场景	1.4
3.核心功能	1.5
3.1 服务端	1.5.1
3.1.0 定时任务	1.5.1.1
3.1.1 全局事务管理	1.5.1.2
3.1.2 全局事务协调	1.5.1.3
3.1.3 异常上报	1.5.1.4
3.1.4 消息上报	1.5.1.5
3.2 客户端	1.5.2
3.2.0 业务拦截	1.5.2.1
3.2.1 全局事务关联	1.5.2.2
3.2.2 事件上报	1.5.2.3
3.2.3 负载均衡	1.5.2.4
3.3 全局事务	1.5.3
3.3.0 手动补偿	1.5.3.1
3.3.1 自动回滚	1.5.3.2
3.3.2 重试	1.5.3.3
3.3.3 超时	1.5.3.4
3.3.4 嵌套	1.5.3.5
3.4 服务降级	1.5.4
3.5 监控	1.5.5
4.配置说明	1.6
4.0 启动前配置	1.6.1
4.1 运行中配置	1.6.2
5.常见问题	1.7
6.产品优势	1.8
7.示例工程	1.9
7.0 Spring Boot示例工程	1.9.1
7.1 Spring Cloud示例工程	1.9.2
7.2 Dubbo示例工程	1.9.3

txle 中文技术参考手册

目录

参考 [gitbook](#) 左侧目录区 或 [SUMMARY.md](#)

PDF下载

[《txle中文技术参考手册.pdf》](#)

官方技术支持

- 代码库 [github](https://github.com/actiontech/txle): github.com/actiontech/txle
- 文档库 [github](https://github.com/actiontech/txle-docs-cn): github.com/actiontech/txle-docs-cn
- 文档库 [github pages](https://github.io/bxle-docs-cn): [actiontech.github.io/bxle-docs-cn](https://github.io/bxle-docs-cn)
- 网站: [TXLE官方网站](#)
- QQ group: 696990638
- 开源社区微信公众号



注意

本分支上的手册适用于todo版txle，其他版本的文档请参考对应tag分支或者release版文档。

联系我们

如果想获得txle 的商业支持, 您可以联系我们:

- 全国支持: 400-820-6580
- 华北地区: 86-13718877200, 王先生
- 华南地区: 86-18503063188, 曹先生

- 华东地区：86-18930110869, 梁先生
- 西南地区：86-18328335660, 雷先生

0.1 txle 简介与整体架构

0.1.1 txle简介

txle是[上海爱可生信息技术股份有限公司](#)基于开源项目[ServiceComb Pack](#)研发的分布式事务开源框架，核心功能是保证业务数据的最终一致性，存在以下几个优势特性：

- **数据最终一致性** 本框架不仅继承了ServiceComb Pack的手动补偿机制，还增设了自动补偿机制，方便业务通过多种方式实现数据最终一致性。
- **高性能** 各子业务均直接执行数据事务提交，而非像XA、TCC等需要锁定数据源。
- **低侵入** 最少情况两个注解即可。
- **支持重试和超时机制** 部分相关框架未能支持重试或超时。
- **支持异常快照处理** 当子事务发生异常时，且未能成功补偿情况下，系统会自动收集相应的异常信息，生成快照，供后续排查分析（如存储本地数据库或上报至差错平台）。
- **支持集成主流框架** 如Spring Boot、Spring Cloud和Dubbo。
- **完善的监控中心** 系统在许多层面进行了重要指标的监控，如服务端、客户端，甚至业务端。
- **最大化保证主营业务的正常运行** 系统提供全局事务的整体与部分降级功能，以及降级后的恢复功能。

0.1.2 txle由来

- txle是基于ServiceComb Pack进行研发的。在此对Saga项目的相关工作人员与贡献者致以由衷的感谢。
- 随着技术的快速发展，微服务架构被广泛地应用到许多系统中。遗憾的是，微服务架构中的事务问题至今还没有很好地落地性的解决方案。txle的出现正是解决微服务架构中的数据最终一致性问题。
- 对公司而言，常年专注于数据库领域相关服务。有关数据库方面，公司已研发出多款产品，如集群高可用(DMP)、分片(dble)、CDC工具(dtile)等相关产品。在数据库事务领域，我们也希望能为业内贡献出我们的绵薄之力。

0.1.3 txle内部架构



3.0 核心功能

- 3.1 服务端功能
 - 3.1.0 定时任务
 - 3.1.1 全局事务管理
 - 3.1.2 全局事务协调
 - 3.1.3 RESTful API
 - 3.1.4 异常上报
 - 3.1.5 消息上报
- 3.2 客户端
 - 3.2.0 业务拦截
 - 3.2.1 全局事务关联
 - 3.2.2 事件上报
 - 3.2.3 负载均衡
- 3.3 全局事务
 - 3.3.0 手动补偿
 - 3.3.1 自动回滚
 - 3.3.3 重试
 - 3.3.4 超时
 - 3.3.5 嵌套
- 3.4 服务降级
- 3.5 监控

服务端定时任务扫描器

- 更新超时状态

查询出【已完成但未标记DONE状态的】超时记录；

更新超时记录为完成状态；

- 查找超时事件

查询(未完成的)超时事件；

保存超时记录；

- 终止超时时间

查询【NEW状态】且【已超时的】超时记录，未超时的超时记录将不做处理；

更新【已超时的】超时记录为【PENDING】状态；

- 查找需补偿事件，保存未补偿命令

查询需补偿事件(即超时异常的和普通异常的)按照从左到右的顺序，优先排除过滤较多的条件；从未结束(即无【SagaEndedEvent】状态)的最小事件开始查询；含发生异常的；排除已补偿的；排除重试的；依据唯一标识升序；

保存未补偿命令；

- 执行补偿

查找状态为NEW的Command；

更新状态为NEW的Command的状态为PENDING；

调用客户端的补偿方法；

- 更新已补偿命令

查找已补偿但未结束的事件；

更新补偿状态为DONE；

事务管理

- 对全局事务及子事务进行新增、修改、删除和查询的相关基本操作。
- 支持子事务重试操作。
- 支持全局事务或子事务的超时。
- 全局事务或子事务在程序中以事件的形式存在。

事务协调

- 当全局事务执行失败时，事务协调器(txle server)会主动对该全局事务内的已完成子事务下达补偿命令，同时对该子事务标记为已补偿。
- 事务协调包含手动补偿和自动回滚。
- 若补偿或回滚失败，被协调客户端会收集异常信息，上报至差错平台并保存至数据库。

上报异常至差错处理平台

- 简介

分布式事务主要保证数据的一致性，如果业务出现异常情况，则会回滚掉所有执行过的子事务。然而，一些异常是系统无法抗拒与处理的(如数据被篡改)，会导致全局事务回滚失败，此时系统会将事务信息上报至差错平台，供相关人员查阅及制定处理方案。

- 配置

```
application.yml
```

```
txle:
  accident:
    platform:
      address:
        api: http://ip:port/receiveFailedGlobalTxInfo
```

```
application.properties
```

```
txle.accident.platform.address.api=http://ip:port/receiveFailedGlobalTxInfo
```

重试次数

```
application.yml
```

```
txle:
  accident:
    platform:
      retry:
        retries: 3
```

重试间隔(秒)

```
application.yml
```

```
txle:
  accident:
    platform:
      retry:
        interval: 3
```

- 上报时机

自动补偿失败时，系统将上报相关信息至差错平台。

手动补偿失败时，由于手动补偿接口由业务人员自行开发维护，故业务人员在编写手动补偿接口时需自行追加上报至差错平台相关代码，参考如下。

如图：

```

application.yml

@.Autowired
private ClientAccidentHandlingService clientAccidentHandlingService;

@Autowired(required = false)
private OmegaContext omegaContext;

public xx xxx_rollback(..) {
    try {
        // 补偿接口的业务代码....
    } catch (Exception e) {
        // 补偿接口失败时，上报差错平台
        JsonObject jsonParams = new JsonObject();
        jsonParams.addProperty("type", AccidentHandleType.ROLLBACK_ERROR.toInteger());
        jsonParams.addProperty("globalthxid", omegaContext.globalTxId());
        jsonParams.addProperty("localtxid", omegaContext.localTxId());

        // bizinfo由业务人员自行收集，如此处为JSON格式的实体信息
        MerchantEntity merchantEntity = merchantRepository.findOne(merchantid);
        if (merchantEntity != null) {
            jsonParams.addProperty("bizinfo", merchantEntity.toJsonString());
        } else {
            jsonParams.addProperty("bizinfo", "{\"merchantid\": " + merchantid + "}");
        }

        LOG.error("Failed to execute method 'updateBalanceById_rollback', params [{}].", jsonParams.toString(), e);
        // 调用上报至差错平台的接口
        clientAccidentHandlingService.reportMsgToAccidentPlatform(jsonParams.toString());
        throw e;
    }
}
}

```

- 监控指标

- 上报成功数量指标 - txle_report_accident_successful_total

记录成功上报至差错平台的数量。

- 上报失败数量指标 - txle_report_accident_failed_total

记录上报至差错平台失败的数量。

- 告警

每次抓取到新的失败记录，将触发告警。

告警规则配置，在alert.rules文件中追加如下代码：

上报失败告警规则

```

# Alert for txle instance that metric txle_report_accident_failed_total value > 0
- alert: ReportAccidentUnsuccessfully(init)
  expr: sum(count_over_time(txle_report_accident_failed_total[1m]) + txle_report_accident_failed_total) == 2
  for: 0s
  annotations:
    summary: "Failed to report accident on {{ $labels.instance }}."

# Alert for txle instance that metric txle_report_accident_failed_total value > 0
- alert: ReportAccidentUnsuccessfully
  expr: idelta(txle_report_accident_failed_total[1m]) > 0
  for: 0s
  annotations:
    summary: "Failed to report accident on {{ $labels.instance }}."

```

上报消息至Kafka平台

- 简介

因子业务无法确定全局事务是否成功完成，即数据是否最终达到一致性，故各子事务不能直接去刷新缓存数据。因此，需等待全局事务完成(无论成功失败)后，将全局事务内部所有子事务执行后影响的数据信息上报到消息管理平台(即 Kafka)，由各子事务自行订阅刷新缓存。

- 消息数据表

名	类型	长度	小数点	不是 null	虚拟	键	注释
id	bigint	20	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		1
globaltxid	varchar	36	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
localtxid	varchar	36	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
status	int	1	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
version	int	2	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		
dbdrivername	varchar	100	0	<input type="checkbox"/>	<input type="checkbox"/>		
dburl	varchar	100	0	<input type="checkbox"/>	<input type="checkbox"/>		
dbusername	varchar	20	0	<input type="checkbox"/>	<input type="checkbox"/>		
tablename	varchar	255	0	<input type="checkbox"/>	<input type="checkbox"/>		
operation	varchar	20	0	<input type="checkbox"/>	<input type="checkbox"/>		
ids	blob	0	0	<input type="checkbox"/>	<input type="checkbox"/>		
createtime	datetime	0	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>		

- 集成逻辑

- 收集数据信息

即数据表Message中有关业务上的字段值信息，如数据库驱动、数据库连接地址、数据库用户名、数据表名称、编辑操作类型、影响数据标识集合。

- 自动补偿场景收集

生成自动补偿信息过程中，收集子事务所影响的业务数据。

- 手动补偿场景收集

因子事务的业务逻辑未知，可能存在数据库方面影响的数据与业务上影响数据不一致，所以应由各子事务自行实现收集所影响的业务数据。

手动补偿情况业务开发人员调用上报Kafka消息的示例代码：

```

@.Autowired
private MessageSender messageSender;

public ... rollback... {
    // 业务代码.....

    // 手动补偿场景：由业务人员自行收集所影响的数据信息
    messageSender.reportMessageToServer(new KafkaMessage(...));
}

```

- 调用统一上报接口

```

org.apache.servicecomb.saga.omega.transaction.MessageSender#
reportMessageToServer(org.apache.servicecomb.saga.omega.transaction.KafkaMessage message);

```

- 上报逻辑

采用gRPC方式进行上报

新增报文消息的proto文件GrpcMessage.proto

逻辑与客户端传输Event到服务端一致，并与其共用同一Endpoint

- TXLE服务端存储业务信息至数据库

即将客户端收集到的业务影响数据保存到数据表Message中。

- TXLE服务端发送业务信息至Kafka服务端

当全局事务结束后，读取数据表Message中相关消息列表，将整个全局事务的消息列表一次性发送至Kafka服务端。

- 发送成功

更新数据表Message中的消息状态status为发送成功。

记录日志。

- 发送失败

更新数据表Message中的消息状态status为发送失败。

记录日志。

上报差错平台。

- 配置文件与热加载

配置文件：..\\txle\\alpha\\alpha-server\\src\\main\\resources\\kafka.properties。

ps: 在测试与生产环境使用前，请先依据当时环境对此文件进行调整，如bootstrap.servers、topic、acks等。

热加载可调用RESTful Api接口：<http://ip:port/reloadConfig/kafka>。

业务拦截

业务发起拦截

- 业务系统集成txle的关键一步是业务发起方法上添加注解@SagaStart，代表该方法使用全局事务。
- txle将会对带有@SagaStart注解的所有方法进行拦截。
- 拦截业务后，开启全局事务，并将其交由事务管理器进行管理
- 接着，执行业务发起函数中的相关子业务程序。

ps：开关全局事务/子事务均需交由事务管理器进行管理。

子业务拦截

- 业务系统集成txle关键的第二步是在子业务方法上添加注解@Compensable或@AutoCompensable，代表该子业务方法使用全局事务的手动补偿或自动回滚功能。
- txle将会对带有@Compensable或@AutoCompensable注解的所有方法进行拦截。
- 拦截子业务方法后，开启全局事务下的子事务，并与当前全局事务进行关联。
- 接着，执行子业务中的相关程序(如操作数据库等)。
- 每个子业务执行完成后，需关闭当前子业务对应的子事务。
- 待所有子业务均执行完成后，关闭全局事务，整个全局事务视为完成。

SQL拦截

全局事务除了对业务方法进行拦截，还会对业务执行的SQL进行拦截，但不会影响业务SQL的正常执行，主要是为了实现自动回滚机制，以及统计SQL的耗时，便于系统性能排查。

全局事务与各子事务的关联

- 全局事务与各子事务是通过全局事务标识xid进行关联的。
- 全局事务发起时会创建当前全局事务的上下文信息，含xid。
- 调用子事务时，将xid隐藏传输至子事务中，实现关联。

可参考调用链路时序图：



上报事件

业务集成全局事务后，全局事务客户端的主要工作就是负责拦截与收集业务信息，随后形成对应的事件，通过RPC传输至服务端。

事件形成时机：

全局事务启动事件：业务发起时产生

子事务启动事件：执行被全局事务管理的子业务时产生

子事务结束事件：被全局事务管理的子业务执行完成后产生

异常中止事件：全局事务内出现异常或超时情况产生

补偿事件：全局事务内出现异常或超时，对子事务进行补偿时产生

全局事务结束事件：业务结束后产生

负载均衡

为防止大量并发集中到某一台服务器上而出现性能下降甚至服务器宕机等隐患，全局事务提供负载均衡功能。

服务端的集群节点支持无限水平扩展，客户端支持配置多个服务端地址，以将请求分散到多台服务。

目前采用的负载均衡算法是最快响应时间算法，即每一次请求完成后会记录响应时间，哪台服务器的响应时间短，则下次就请求到就使用哪台服务器。

PS：目前算法存在一定的缺陷，即某一时间段大量并发请求可能会都被请求到上次响应时间最快的服务器上，该问题正在处理中，待处理后会更新此文档。

手动补偿

手动补偿功能是保证(微服务)分布式事务场景下的数据最终一致性的核心功能。

手动补偿功能需通过在业务方法上添加注解@Compensable和编写对应的补偿接口来启用，补偿接口需被编写在同类中，且接口返回类型、参数类型等均要与被补偿的接口保持统一。

若开启了手动补偿功能，则子业务执行前开启子事务的同时会记录对应的补偿方法到当前子事务事件中，供后续(全局事务出现异常时)补偿调用。

手动补偿失败时，会收集全局事务和业务的异常信息并形成异常快照[上报至差错平台](#)。

手动补偿接口需支持幂等性。

自动回滚

自动回滚机制同手动补偿实现思想几乎一致，主要区别是自动回滚通过对业务SQL的代理完成自动识别业务SQL功能，并生成反向补偿的对应SQL，在后续需要补偿时执行反向SQL。

客户端通过增设回滚日志数据表来记录相应事务的反向SQL。

重试机制

全局事务针对各子事务支持重试功能，以避免因网络抖动等因素影响到业务的正常完成。

开启/关闭方法

通过在注解@Compensable或@AutoCompensable中设置retries来开启或关闭重试功能。

设置该参数值设置为0或不设置该参数，代表不使用重试功能。

设置大于0的正整数参数值，代表启用重试功能。

参数值

- 默认未设置该参数，即值为0，代表不重试。
- 当值为-1时，无限重试。
- 超时情况，暂不支持重试。

间隔时间

重试间隔时间默认为3s，可通过参数[omega.connection.reconnectDelay]进行配置。

注意事项

不支持超时场景下的重试。

补偿接口不支持重试。

重试信息修改后需重启应用才能生效。

超时机制

全局事务与各子事务均支持超时功能。

开启/关闭方法

通过在注解@Compensable或@AutoCompensable中设置timeout来开启或关闭超时功能。

参数值

默认未设置该参数，即值为0，代表不使用超时功能。

设置大于0的正整数值，代表启用超时功能。

★★★超时检测★★★

后台定时任务检测

在全局事务正常运行过程中，定时任务会每隔0.5s对全局事务进行超时检测。

全局事务结束前检测

部分场景，定时任务执行时间大于全局事务的超时时间，这导致了还未等对全局事务进行超时检测，结果全局事务就已经结束了。为杜绝此现象，增设全局事务结束前检测超时。

超时处理

检测到超时的全局事务后，先后创建超时记录、创建中止事件、创建待补偿命令、下发补偿命令、创建补偿事件、更新超时记录状态、更新待补偿命令、场景全局事务完成事件。

超时场景会对超时的全局事务下的**所有子事务都进行补偿**，而非超时异常仅针对当前全局事务下的非当前异常的**其它子事务进行补偿**。

注意事项

超时信息修改后需重启应用才能生效。

事务嵌套

- 嵌套全局 事务

即有两个 定义为全局 事务(@SagaStart)接口 A、B，A中调用 B，当 B 发生错误时，A、B 中的操作都应回滚。

结果：ok。

- 手动补偿子事务嵌套

嵌套 Compensable 事务，即 Saga 中有补偿方法 A、补偿方法 AA，A中调用 AA。程序执行后，仅记录一个 tx 事件，说明 AA未被看做为一个单独的补偿方法。

- 自动补偿子事务嵌套同手动补偿子事务嵌套

全局事务服务降级

- 简介

大部分框架和中间件在设计初期均应考虑后续对主营业务的影响最小化，txle严格遵守该原则，提供服务降级功能。在程序运行时的多个阶段都可对全局事务进行降级或半降级，以及降级后的恢复功能，这在很大程度上保证了业务的正常运行。

- 背景

性能影响：当全局事务严重影响业务系统性能，且业务数据一致性要求不严苛时，可对全局事务采取降级处理。

Bug：当全局事务运行过程中出现bug时，未能保证业务数据不一致时，如未能及时定位问题与修复，可对全局事务采取降级处理。

异常：当全局事务运行过程中出现异常时，会回滚所涵盖的子业务功能，导致业务无法正常进行，可对全局事务采取降级处理。

特殊情况：除全局事务外，系统内其他主要功能，如监控、Kafka集成、差错平台等功能也提供了降级处理，主要作用都是在特殊情况不影响业务的正常运行。

- 配置中心数据表

名	类型	长度	小数点	不是 null	虚拟	键
id	bigint	20	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	1
servicename	varchar	100	0	<input type="checkbox"/>	<input type="checkbox"/>	
instanceid	varchar	100	0	<input type="checkbox"/>	<input type="checkbox"/>	
category	varchar	100	0	<input type="checkbox"/>	<input type="checkbox"/>	
type	int	2	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
status	int	1	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
ability	int	1	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
value	varchar	100	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
remark	varchar	500	0	<input type="checkbox"/>	<input type="checkbox"/>	
updatetime	datetime	0	0	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

```

CREATE TABLE IF NOT EXISTS Config (
    id bigint NOT NULL AUTO_INCREMENT,
    servicename varchar(100),
    instanceid varchar(100),
    type int(2) NOT NULL DEFAULT 0 COMMENT '1-globaltx, 2-compensation, 3-autocompensation, 4-bizinfotokafka, 5-txmonitor, 6-alert, 7-schedule, 8-globaltxfaulttolerant, 9-compensationfaulttolerant, 10-autocompensationfaulttolerant, 50-accidentreport, 51-sqlmonitor if values are less than 50, then configs for server, otherwise configs for client.',
    status int(1) NOT NULL DEFAULT 0 COMMENT '0-normal, 1-historical, 2-dumped',
    ability int(1) NOT NULL DEFAULT 1 COMMENT '0-do not provide ability, 1-provide ability      ps: the client''s ability inherits the global ability.',
    value varchar(100) NOT NULL,
    remark varchar(500),
    updatetime datetime NOT NULL DEFAULT CURRENT_TIMESTAMP,
    PRIMARY KEY (id) USING BTREE,
    UNIQUE INDEX pk_id(id) USING BTREE,
    INDEX index_type(type) USING BTREE
) DEFAULT CHARSET=utf8;

```

- 配置管理

理应参考配置中心相关文档（暂未开发）。

提供展示UI，读取配置表数据并展示。

添加时，客户端名称和实例ID从全局事件数据表中读取。

由于数据量不大，故更新时先将原数据置为历史数据，再插入新数据。

删除全部为逻辑删除，即删除后状态变为历史数据。

- 降级类型

- 配置降级

即通过配置对部分功能进行降级处理与否。

系统提供部分核心功能的降级配置，当功能对应的配置值被设置为disabled时，则说明此功能被设置为降级处理，即程序暂停提供此功能直至解除降级。若程序运行一切正常，可通过设置配置值为enabled解除降级。

当然，降级配置值不仅仅包含disabled，后续会增设其它值，如范围值等。

- 容错降级

即全局事务系统运行过程中发生异常时，为不影响业务正常运行，将当前异常捕获掉，不回滚子业务已提交的事务。

容错降级也支持配置，即如果启用了相应功能的容错降级才会对异常进行捕获。

- 降级配置

服务器端降级配置：全局事务、手动补偿、自动补偿、Kafka集成、事务监控、告警、定时器。

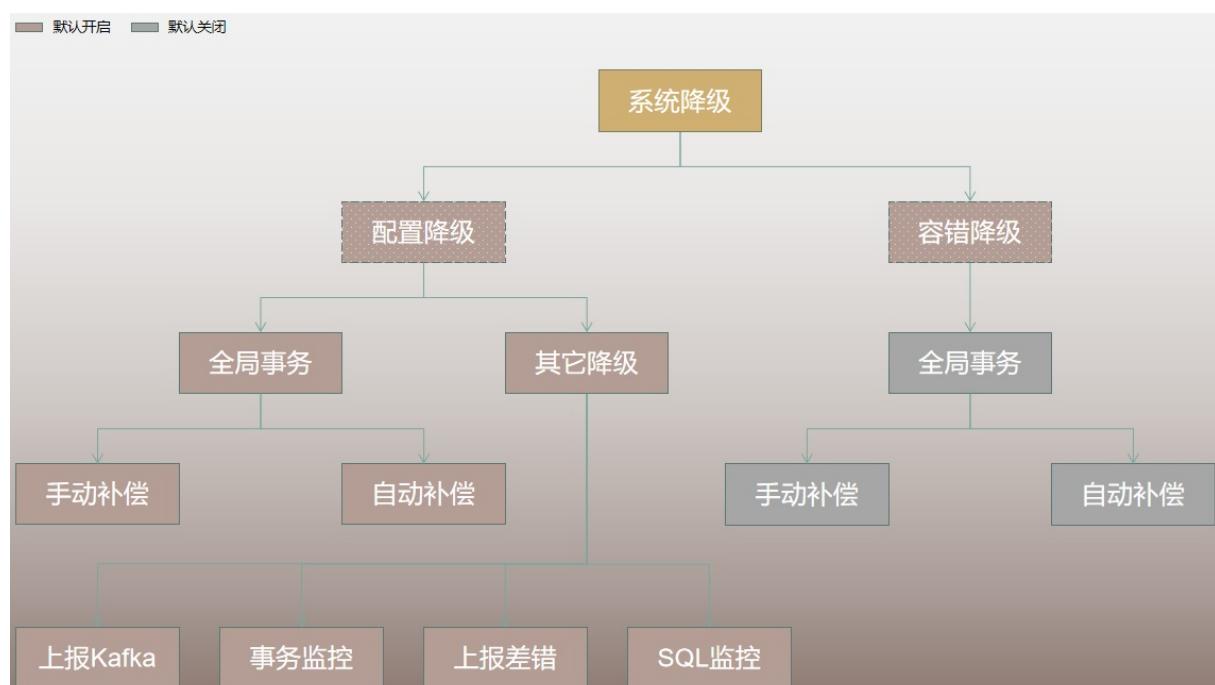
容错降级配置：全局事务容错、手动补偿容错、自动补偿容错。

客户端降级配置：上报差错、SQL监控。

ps：降级是概念上的，而真实的系统配置是针对实际的功能十分支持。

如针对降级配置，开启全局事务，意思是业务系统使用全局事务，关闭则是业务系统不使用全局事务；

如针对容错降级，开启全局事务是发生错误时，将进行容错，即不对子事务进行回滚，关闭时，如果发送错误，则会对子事务进行回滚操作；



全局事务：

在全局事务启动前会先检测是否启用全局事务配置，如配置值为disabled则不启动全局事务，直接执行业务。

每个全局事务启动前都会检查，以达到降级配置实时生效。已在运行的不受影响。

手动补偿： 同上。在启动子事务前检测。

自动补偿： 同上。

Kafka集成： 同上。在发送Kafka消息前检测。

事务监控： 在启用监控前检测。

告警： 暂未开发，可在Prometheus的配置文件中进行配置启用/禁用。

定时器： 暂未开发，目前已有定时器均针对所有客户端，故不便于降级。

全局事务容错： 在全局事务启动至结束代码处的异常捕获中增设容错降级功能，依据是否启用全局事务容错降级配置进行容错降级处理。

手动补偿容错： 同上。在手动补偿对应的子事务代码处。

自动补偿容错： 同上。在自动补偿对应的子事务代码处。

上报差错： 发生消息到差错平台前检测，依据该功能的降级配置值决定是否继续上报。

SQL监控： 客户端业务执行过程中，如果开启全局事物则在手动补偿或自动补偿子事务启动前检测当前配置值，如果未开启全局事物则在业务执行前检测。

- 降级逻辑

系统针对部分核心功能提供降级配置。

默认除容错降级配置为disabled，其它降级配置值均为enabled。

每项配置都分为全局配置和客户端配置，全局配置提供是否支持某功能。若不支持，则采用全局默认值；若支持，优先客户端配置，客户端未配置则采用全局默认值；私有配置值不能超出全局配置值范围。

相关功能执行前，依据其对应的降级配置决定是否继续执行该功能。

由于降级需求往往都比较急迫，故所有配置均提供实时生效功能。

各配置检测时机不同，参考设计原则是“保证实时失效，节省性能”。

监控与告警

简介

监控主要对系统中的一些重要指标进行采集，采集后可友好地展示在UI中，便于后续对系统的性能、吞吐、稳定性等整体分析。

告警是依据采集的指标配合相关公式设计的告警规则，可及时地发现系统以及服务器存在的问题，避免时间长未处理而酿成重大后果。

除本系统重要指标外，也包括JVM、CPU、内存、磁盘和网络等。

txle监控指标

TXLE项目

全局事务数量

- 成功数量 - `txle_transaction_total`
- 失败数量 - `txle_transaction_failed_total`
- 回滚数量 - `txle_transaction_rollbacked_total`
- 超时数量 - `txle_transaction_timeout_total`
- 重试数量 - `txle_transaction_retried_total`
- 暂停数量 - `txle_transaction_paused_total`
- 继续数量 - `txle_transaction_continued_total`
- 自动继续数量 - `txle_transaction_autocontinued_total`

采用Prometheus的Counter类型统计，部分代码如下：

```
private final Counter txleTransactionTotal = buildCounter( name: "txle_transaction_total", help: "Total number of transactions.");
private final Counter txleTransactionSuccessfulTotal = buildCounter( name: "txle_transaction_successful_total", help: "Total number of successful transactions.");
private final Counter txleTransactionRollbackedTotal = buildCounter( name: "txle_transaction_rollbacked_total", help: "Total number of rollbacked transactions.");
private final Counter txleTransactionFailedTotal = buildCounter( name: "txle_transaction_failed_total", help: "Total number of transactions which had abnormality occurs.");
private final Counter txleTransactionRetriedTotal = buildCounter( name: "txle_transaction_retried_total", help: "Total number of retried transactions..");
private final Counter txleTransactionTimeoutTotal = buildCounter( name: "txle_transaction_timeout_total", help: "Total number of timeout transactions.");
private final Counter txleTransactionPausedTotal = buildCounter( name: "txle_transaction_paused_total", help: "Total number of paused transactions.");
private final Counter txleTransactionContinuedTotal = buildCounter( name: "txle_transaction_continued_total", help: "Total number of continued transactions.");
private final Counter txleTransactionAutocontinuedTotal = buildCounter( name: "txle_transaction_autocontinued_total", help: "Total number of auto-continued transactions.");

private Counter buildCounter(String name, String help) {
    return Counter.build(name, help).labelNames("business", "category").register();
}
```

```

public void countTxNumber(TxEvent event, boolean isTimeout, boolean isRetried) {
    if (!isEnableMonitor(event)) {
        return;
    }
    try {
        Set<String> eventTypesOfCurrentTx = globalTxIdAndTypes.get(event.globalTxId());
        if (eventTypesOfCurrentTx == null) {
            eventTypesOfCurrentTx = new HashSet<>();
        }
        if (globalTxIdAndTypes.isEmpty()) {
            // release memory
            globalTxIdAndTypes.clear();
        }
        String type = event.type(), serviceName = event.serviceName(), category = event.category();
        if (!eventTypesOfCurrentTx.contains(type)) {
            eventTypesOfCurrentTx.add(type);
            globalTxIdAndTypes.put(event.globalTxId(), eventTypesOfCurrentTx);

            if (SagaStartedEvent.name().equals(type)) {
                // ps: it would not appear in the metrics page if didn't set the labels' values.
                txleTransactionTotal.labels(serviceName, category).inc();
            } else if (SagaEndedEvent.name().equals(type)) {
                txleTransactionSuccessfulTotal.labels(serviceName, category).inc();
                globalTxIdAndTypes.remove(event.globalTxId());
                return;
            } else if (TxAbortedEvent.name().equals(type)) {
                txleTransactionFailedTotal.labels(serviceName, category).inc();
            } else if (TxCompensatedEvent.name().equals(type)) {
                txleTransactionRollbackedTotal.labels(serviceName, category).inc();
            } else if (AdditionalEventType.SagaPausedEvent.name().equals(type)) {
                txleTransactionPausedTotal.labels(serviceName, category).inc();
            }
        }
    }
}

```

子事务数量

指标名称：txle_transaction_child_total

采用Prometheus的Counter类型统计，部分代码如下：

```

public void countChildTxNumber(TxEvent event) {
    if (!isEnableMonitor(event)) {
        return;
    }
    if (TxStartedEvent.name().equals(event.type()) && !localTxIdSet.contains(event.localTxId())) {
        txleTransactionChildTotal.labels(event.serviceName(), event.category()).inc();
        if (event.retries() > 0) {
            localTxIdSet.add(event.localTxId());
        }
    } else if (TxEndedEvent.name().equals(event.type()) || SagaEndedEvent.name().equals(event.type()) || TxAbortedEvent.name().equals(event.type())) {
        localTxIdSet.remove(event.localTxId());
    }

    if (localTxIdSet.isEmpty()) {
        log.info("The 'localTxIdSet' variable is empty, and it will be collected when JVM executes GC at the next time.");
        localTxIdSet.clear();
    }
}

```

全局事务时长

指标名称：txle_transaction_time_seconds_total

```
// Total seconds spent for someone transaction. Ps: It will show one row only if you search this metric in 'http://ip:9090/graph'.
// But, Prometheus will record every metric in different times, and you can search, like 'txle_transaction_time_seconds_total[5m]', then it will show many rows to you.
private final Gauge txleTransactionTimeSecondsTotal = buildGauge( name: "txle_transaction_time_seconds_total", help: "Total seconds spent executing the global transaction.");
private final Gauge txleTransactionChildTimeSecondsTotal = buildGauge( name: "txle_transaction_child_time_seconds_total", help: "Total seconds spent executing the child transaction.");

private Gauge buildGauge(String name, String help) {
    return Gauge.build(name, help).labelNames("business", "category").register();
}

public void startMarkTxDuration(TxEvent event) {
    if (!isEnabledMonitor(event)) {
        return;
    }
    // Start a timer to track a duration, for the gauge with no labels. So, we must set the value of the labelNames property.
    if (SagaStartedEvent.name().equals(event.type())) {
        txIdAndGaugeTimer.put(event.globalTxId(), txleTransactionTimeSecondsTotal.labels(event.serviceName(), event.category()).startTimer());
    } else if (TxStartedEvent.name().equals(event.type())) {
        txIdAndGaugeTimer.put(event.localTxId(), txleTransactionChildTimeSecondsTotal.labels(event.serviceName(), event.category()).startTimer());
    }
}
```

子事务时长

指标名称：txle_transaction_child_time_seconds_total

同全局事务时长，将globalTxId改为localTxId。

业务SQL时长与数量

指标名称：txle_sql_time_seconds_total、txle_sql_total

即业务应用系统执行的SQL。

标准业务SQL统计时已依据业务和事务类别分组。

未分组的业务SQL主要包含：自动补偿SQL。

注：当某子业务被成功回滚时，对应的txle_sql_total会有7条，即针对当前子事务会有如下操作“开启事务、结束事务、查询回滚状态、新增待完成的回滚命令、新增回滚事件、更新回滚命令为完成、更新全局事务结束事件”。

维护SQL时长与数量

指标名称：txle_sql_time_seconds_total、txle_sql_total

即TXLE项目执行的SQL，含Client端自动补偿相关维护SQL和Server端的维护(或所有)SQL。

标准维护SQL统计时已依据业务和事务类别分组。

未分组的维护SQL主要包含：Client端自动补偿SQL、Server端部分定时器SQL(无全局事务标识参数的)、超时和补偿事件SQL。

JVM

采用prometheus自带统计方式：

```
io.prometheus.client.hotspot.DefaultExports.initialize();
```

CPU、内存、磁盘、网络

安装第三方工具**Node exporter**来统计指标。

TXLE默认Prometheus客户端服务端口

TXLE Server端默认Prometheus端口8099

默认端口逻辑：默认端口配置在**application.yaml**文件中，所以程序中获取到的任何值均可认为是维护人员主动编辑，所以在程序层面如果获取到无效值，则意味着不启动Prometheus客户端监控服务。

application.yaml

```
txle:  
  prometheus:  
    metrics:  
      port: 8099
```

TXLE Client端默认Prometheus端口8098

默认端口逻辑：Client与Server不一样，Client端的默认端口只能写在**Java代码**中。因为Client端的Prometheus服务端口需由集成业务进行设置，所以无法在Client端进行默认配置。

application.yaml

```
txle:  
  prometheus:  
    metrics:  
      port: 8098
```

application.properties

```
txle.prometheus.metrics.port=8098
```

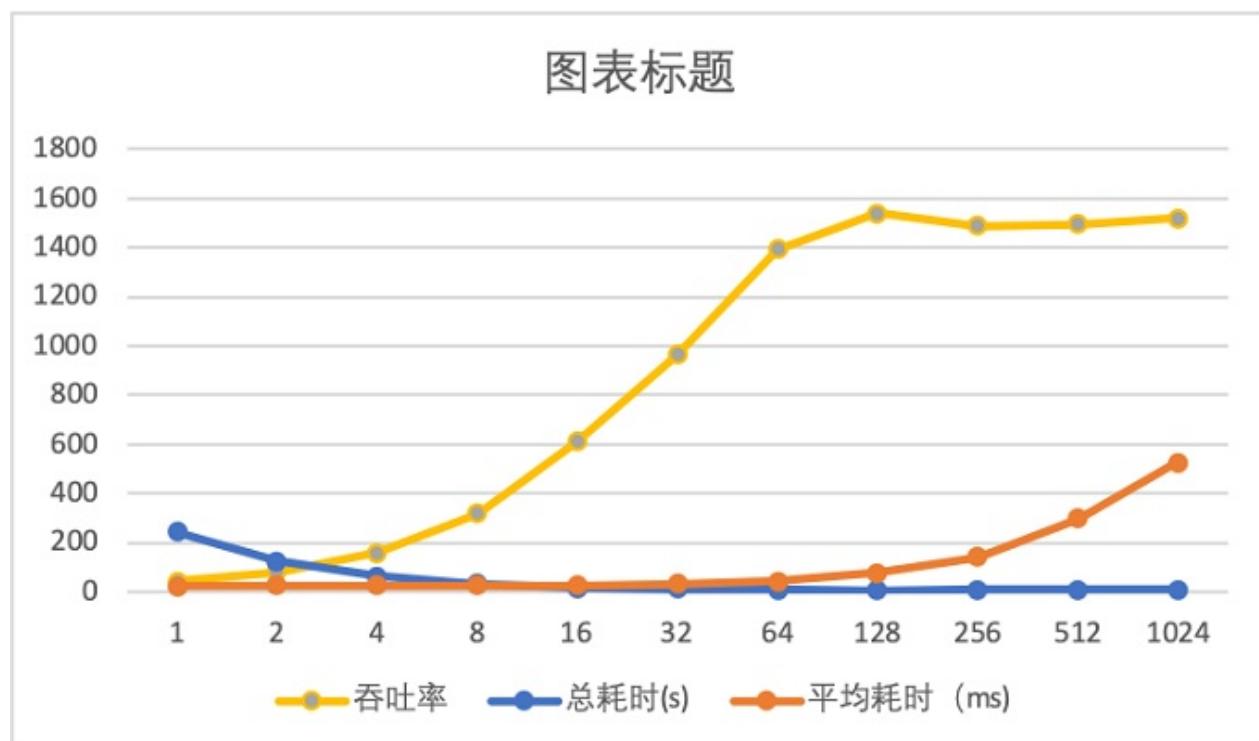

常见问题

是否支持docker快速部署？

支持常规部署、docker镜像部署、以及docker compose部署。

QPS/TPS?

tx|e QPS 指标：



ServiceComb Pack QPS指标：

节点数	并发数	总耗时(s)	平均耗时(ms)	吞吐率	事务量
0.5.0 without Akka	100	102	1026	98/sec	10000
0.5.0 without Akka	500	99	4970	101/sec	10000
0.5.0 with Akka	100	14	142	714/sec	10000
0.5.0 with Akka	500	8	418	1250/sec	10000
0.5.0 with Akka	1000	8	858	1250/sec	10000
0.5.0 with Akka	2000	14	2888	714/sec	10000
0.5.0 with Akka	1000	39	786	1282/sec	50000
0.5.0 with Akka	2000	37	1519	1351/sec	50000
0.5.0 with Akka	3000	43	2687	1116/sec	50000

保证ACID么？

保证ACD，不保证隔离性。原因是产品的核心主要解决数据的最终一致性。

支持哪些数据库？

目前我们专注于MySQL，后期计划增加PostgreSQL等。

对数据库连接池有特殊要求么？

暂无特殊要求。

对于补偿方法的要求？

- 需要与被补偿方法在同一类且参数完全一致
- 参数可被序列化

是否支持事务嵌套？

支持子事件嵌套。

txle server端支持水平扩展么？

支持。相关状态信息均存储于数据库。

txle server宕机重启后，是否可继续当前全局事务？

可以继续当前全局事务。

对业务系统的性能影响？

单事务分支平均响应时间在2ms左右。

6.0 txle vs other softwares

6.0.1 txle vs XA/TCC

- **高性能** 无需锁定数据资源
- **多重补偿方式** 支持手动补偿与自动补偿机制

6.0.2 txle vs Seata

- **高性能**
- **手动补偿** 相比Seata自带的自动补偿机制更加灵活，支持复杂业务场景

6.0.3 txle vs ServiceComb Pack

- **高性能** 提升几倍的QPS
- **多重补偿方式** 支持自动补偿机制，相比手动补偿机制，节省了大量的开发成本

6.0.3 txle vs other softwares

- **高性能**
- **专注金融领域**
- **支持服务降级**
- **支持差错处理**

