



Transwarp Community Edition 2024-5-CE 使用指南

星环信息科技（上海）股份有限公司

版本号 T00993x-92-010, 2024-6

目录

1. 社区版家族介绍	2
1.1. 关于社区版您可能想要知道的	2
2. 安装及使用前须知（重要）	4
2.1. 安装及使用注意事项	4
2.1.1. 社区版	4
2.1.2. 社区开发版	6
3. 常见运维指令	9
3.1. 社区版Manager操作指令集	9
3.2. 社区版开发版镜像&容器相关指令	9
3.3. 常见的Linux运维指令	9
3.3.1. 目录操作	9
3.3.2. 文件系统以及权限	10
3.3.3. 文件和目录操作	11
3.3.4. 文本处理，查看文件内容	11
3.3.5. 进程管理	12
3.3.6. 用户和权限管理：	12
3.3.7. 系统信息	12
3.3.8. 网络管理	12
3.3.9. 包管理	12
3.3.10. Linux的链接	13
3.3.11. Vim编辑器基本操作	13
3.3.12. 账号管理	13
3.4. 常见的Docker指令	14
3.4.1. 常用的基础操作	14
3.4.2. 其他进阶操作	15
3.5. 常见的K8S命令	16
4. 安装前可能想要了解的产品问题	18
4.1. TDH社区开发版与TDH社区版能否部署在同一服务器？	18
4.2. 星环科技是否提供 Hive ？	18
4.3. Quark是什么？	18
4.4. ArgoDB/Inceptor/Hyperbase三者有什么区别？	18
4.5. TDH集群可以部署两个HDFS服务和Yarn服务吗？	19
4.6. Manager可以添加多个Quark服务吗？	19
4.7. TOS Registry可以多活部署吗？	19
4.8. 社区开发版想换成社区版，使用同一节点需要什么操作吗？	20
4.9. 如果数据源都是通过API接口方式获取，星环平台有什么工具是可以用作数据采集的吗？	20
4.10. 像上层组件(Hybase, Scope, StellarDB..)这些最终数据的存储，其实也是落地在Inceptor 中，建立相应的数据映射，是否可以这样理解？	20
4.11. 星环可以对外提供数据服务吗，外部程序如何获取数据	21
4.12. 社区版对国产化支持度怎么样？可以在麒麟V10里安装吗？	21
5. 【TDH社区版】安装Manager向导过程中可能遇到的问题	22
5.1. 【RPM配置教程】安装向导显示需要配置RPM库，如何配置？	22

5.2.	进入Manager 8179安装页面显示需升级或安装失败后安装页面变为升级页面	26
5.3.	安装向导检查环境页面报错显示“无法解析hostname”	26
5.4.	安装向导显示提取文件失败，报错“The package container-selinux conflicts with manager…27”	
5.5.	配置RPM库显示失败，报错显示“failed to mount ISO file /opt/xxx.iso”	28
5.6.	安装过程中报错显示“Fail to start database service”	28
5.7.	安装失败，报错显示“The directory /var/lib/docker is not on a separate partition”	29
5.8.	安装失败，报错显示“failed to install ntp”	30
5.9.	安装失败，报错显示“系统已经安装过Manager-8.1.3-1.noarch”，但服务连不上	30
5.10.	安装失败，报错显示“failed to install MariaDB server”	31
6.	配置集群过程中可能遇到的问题	33
6.1.	添加节点过程中报错显示“主机名或IP为xxx的节点不存在”	33
6.1.1.	用户现象1.	33
6.1.2.	用户现象2.	34
6.2.	安装集群xx中的3个节点失败，日志报错显示：error.nodeHostNotFound.	34
6.3.	添加节点步骤报错显示“节点无法加入，原因：error.nodeGetHostnameFail”	36
6.4.	添加节点时，访问节点身份使用sudo用户，报错显示“节点无法加入，sudo:需要密码”	36
7.	安装产品服务过程中可能遇到的问题	38
7.1.	通用的排错方法	38
7.1.1.	添加服务阶段初始化服务失败如何定位问题	38
7.2.	上传产品包失败	45
7.3.	许可证页面打不开、上传产品包失败	46
7.4.	安装过程中，配置Quark失败，报错步骤为“生成Collector配置，配置Quark Metastore”	47
7.5.	安装Transwarp Inceptor的时候，配置Quark Metastore(tdh1)的job报错	48
7.6.	安装Guardian服务报错：“/api/market/install/advices: 500 Internal Server Error”	50
7.7.	安装Guardian服务报错：“/api/services/19:500 Internal Server Error”	51
7.8.	安装guardian的时候报错显示“error code 48000…unauthorized”	51
7.9.	第二次安装Guardian时卡住了，报错显示“error code 403…Access is denied”	52
7.10.	社区开发版安装过程中，初始化HDFS失败.	53
7.11.	Dashboard安装完成后状态显示不可用	54
7.12.	TOS服务显示待安装状态，安装时报错无法添加register角色	56
8.	使用过程中可能遇到的问题	57
8.1.	某个服务突然挂掉了的常规排查步骤	57
8.2.	某个服务启动失败的常规排查步骤	57
8.2.1.	示例1.启动TDH服务失败如何定位问题.	57
8.2.2.	示例2.误操作，社区开发版容器停止运行，如何恢复	61
8.3.	Manager及其他服务出现问题.	62
8.3.1.	Manager相关问题.	62
8.3.2.	Yarn相关问题	64
8.3.3.	Quark相关问题.	65
8.3.4.	Guardian相关问题	69
8.3.5.	HFDS相关问题	77
8.3.6.	TOS相关问题.	78
8.3.7.	Aquila相关问题	78
8.3.8.	其他问题	79

8.4. 端口冲突相关问题	79
8.4.1. 什么是端口?	80
8.4.2. 如何避免出现端口冲突, 保障集群顺利安装以及稳定运行	80
8.4.3. 怎么判断是出现了端口冲突	80
8.4.4. 如何解决端口冲突	81
8.4.5. 其他获取冲突端口信息的方式	81
8.4.6. 端口连通性测试	81
8.4.7. 一些排错示例	82
8.5. 客户端相关问题	84
8.5.1. 点击下载客户端无反应	84
8.5.2. Manager中TDH-Client下载失败.	85
8.5.3. 如何让TDH客户端初始化失效?	85
8.5.4. 更换/新增了服务, 之前下载的客户端是否会冲突?	85
8.5.5. 开启Guardian后, TDH-Client报错, 安装了kerberos client, hdfs也报错.	86
8.6. Waterdrop连接与使用相关问题	86
8.6.1. Waterdrop是什么, 如何安装?	86
8.6.2. Waterdrop下载后, 打开报错显示: The Waterdrop executable launcher was unable to locate its companion shared library	87
8.6.3. Windows系统打开Waterdrop时出现无JRE/JDK报错, 应如何解决?	87
8.6.4. Waterdrop 1.2.x 在MacOS Ventura 13.x上无法正常运行	92
8.6.5. Waterdrop报错 COMPILE FAILED:Semantic error: [Error 10011] Line 1:7 Invalid function. Error encountered near token 'LICENSE'	94
8.6.6. Waterdrop如何获取Keytab文件?	95
8.6.7. waterdrop 突然变卡顿是为什么	97
8.6.8. waterdrop可以设置行号吗?	97
8.6.9. waterdrop-inceptor: filter表如何恢复	98
8.6.10. Waterdrop如何把数据库中表的数据导出到本地?	99
8.6.11. 基于Waterdrop进行数据开发时, SQL Script如何存储以及导入新的SQL脚本?	99
8.6.12. Waterdrop调用SYSDATE函数返回结果没有时分秒	101
8.7. 数据库连接相关问题	103
8.7.1. 如何连接数据库?	103
8.7.2. 请问principal的名称如何获取?	103
8.7.3. 集群外连接不上, 报错“could not open connection to jdbc: xxx”怎么排查.	104
8.7.4. 使用beeline命令连接Inceptor时, 报错“JAVA_HOME /usr/lib/jvm/java-xxx/bin/java does not exist”是为什么	104
8.7.5. 使用beeline命令连接Inceptor时, 报错“-bash: beeline: command not found”	105
8.7.6. 使用beeline命令连接Inceptor时, 报错“GSS initiate failed”	105
8.7.7. 使用beeline连接串进入命令行失败, 报错显示‘Peer indicated failure: Unsupported mechanism type PLAIN (state=,code=0)’	105
8.7.8. 基于Windows如何使用ODBC配置使用inceptor	108
8.7.9. Squirrel SQL连接Inceptor时, 连接时报错“aaaa:Connection is closed! Class java.sql.SQLException:Connection is closed”	108
8.7.10. kerberos登录inceptor数据库, kinit hive的密码多少, 在哪看	109
8.7.11. 开启了kerberos后无法采用beeline的方式无法连接数据库, 报错这是为什么?	110
8.7.12. 开启guardian安全且认证模式是LDAP, 用kerberos连接串无法连接Quark, 为什么?	110

8.8. 语法&数据库使用相关问题	111
8.8.1. 使用数据库时报错怎么解决?	111
8.8.2. orc表删除列时, 报错Replacing columns cannot drop columns for table.	111
8.8.3. 在TDH社区版的Inceptor中, 使用Insert into插入单条数据失败, 请问是什么原因呢?	112
8.8.4. 存储过程跑一个多小时后就自动停止了, 报错 operation cancelled.	112
8.8.5. 并发执行插入sql的话会报移动文件错误, 这个有解决办法吗?.	112
8.8.6. 在跑批中经常出错, 取消后重跑又正常, 报错显示FetchFailed(BlockManagerIdxxx...).	113
8.8.7. SQL中Select后字段加别名可正常运行, 不加则报错	113
8.8.8. 查询sql出现错误: lineLength is longer than buffersize, 是否有解决方案	114
8.8.9. 租户quark的权限都配置了, 但是访问的时候还是提示没权限	114
8.8.10. 用insert overview方法把一个orc表的数据导出生成数据文件, 导出时没有分区, 重新load进去时如何指定分区?	115
8.8.11. Inceptor通过使用HiveMetaStoreClient获取表信息的时候出错, 应该怎么处理呢.	116
8.8.12. 如何获悉索引信息?	117
8.8.13. 如何统计inceptor中某一表的使用频率?	117
8.8.14. 如何获取一个数据库中存储的所有表元数据信息, 包括表名, 字段, 表数据条数, 表大小, 表空间等等?	118
8.8.15. Inceptor可以设置主键自动增长吗?	119
8.8.16. inceptor没有distinct on (字段)去重的语句, 如何创建视图(取最新的)	119
8.8.17. 怎么生成UUID?	120
8.8.18. 怎么计算wordcount?	120
8.8.19. 分桶按多少数据量分比较好?	120
8.8.20. insert into批量插入数据, 会有行数限制吗?	120
8.8.21. altertable怎么添加orc表的字段?	121
8.8.22. 如何查看执行查询的时候有没有走索引	121
8.8.23. 星环大数据平台存储中文占据几个字节?	121
8.8.24. 用日期分区, 只能用范围分区吗?	121
8.8.25. 有什么办法查询分区表某一分区的文件大小么?.	122
8.8.26. 向分区表导入数据的时候, 必须要添加partition(分区字段=? ?)条件吗?	122
8.8.27. 本地load csv 到星环表中报错.	122
8.8.28. 如何将 csv 导入 hdfs 中?	123
8.8.29. 安装完的HDFS用户是什么, 如何切换到hdfs的用户上传文件?.	124
8.8.30. 外部表指定一个文件路径后, 若当前路径下有多个文件, 如何处理?.	124
8.8.31. 加载一个文件, 一直说路径不存在	124
8.8.32. SQL游标没有内容, 循环进不去.	125
8.8.33. 用odbc执行生成函数, sql执行 !set plsqlUseSlash true无法处理.	125
8.8.34. 用oracle方言写函数或存储过程, 中间不允许有分号吗?	125
8.8.35. TDH sql, 在分段 union all时, 字符类型会自动补充空格.	126
8.8.36. 有间隔符的数据源, 如何划分task是多少?	126
8.8.37. TDH-insert overwrite local ...row format delimited_fields terminated by ' @ '分隔符问题	127
8.8.38. 正在把oracle的程序迁到星环平台, 数据排序的顺序是怎么样的?	127
8.8.39. 表有重复数据发生, 能否使用rowkey来确保唯一?	128

8.8.40.	函数CURRENT_DATABASE是否可以省略from system.dual?	128
8.8.41.	请问argodb或inceptor的物化视图日终刷新是怎样的?	128
8.8.42.	杀任务, 能通过beeline执行吗?	128
8.8.43.	基于JDBC对Hyperbase数据库进行数据开发时, 手册中配置文件如何获取?	129
8.8.44.	Hyperbase 通过 JAVA API 进行访问, 开启安全后如何配置呢?	129
8.8.45.	是否有最新的Hyperbase API Demo	130
8.8.46.	hbase:acl表不存在.	130
8.9.	小文件以及合并操作的相关问题	131
8.9.1.	小文件背景介绍、不同阶段解决方案、不同表格式合并方式等相关链接总览.	131
8.9.2.	当出现小文件相关的问题以及应该怎么处理.	131
8.9.3.	TORC Compact 常见故障诊断	132
8.9.4.	Compaction Blacklist机制.	134
8.9.5.	Insert into产生大量小文件后采用Compact机制不生效, 无法提交对应的任务到执行引擎中	135
8.10.	自定义函数UDF、UDAF、UDTF相关问题	137
8.10.1.	如何使用外部jar包中自定义函数UDF、UDAF、UDTF	137
8.10.2.	如何创建持久化函数?	137
8.10.3.	UDF自定义函数一次支持多少数据使用.	138
8.10.4.	UDF函数调用时报错(一)	138
8.10.5.	UDF调用报错(二)	138
8.10.6.	星环UDTF函数的输出相关的问题	140
8.11.	数据倾斜引起的问题	140
8.11.1.	星环hive表的数据倾斜导致查询跑不动了怎么解决	140
8.12.	其他产品相关问题	141
8.12.1.	Argodb分页能否用limit?	141
8.12.2.	ArgoDB支持对字符串进行SM3加密吗, 有支持的函数吗?	141
8.12.3.	请问argodb如果获得表的索引信息?	141
8.12.4.	Argodb中, 取最新8位日期时, REPLACE和TO_CHAR的返回结果存在BUG, 求解?	141
8.12.5.	ArgoDB创建好的存储过程的创建语句怎么查?	143
8.12.6.	ArgoDB 写入 HoloDesk 表报错, 超过 2500 万之后数据无法写入.	143
8.12.7.	星环哪个组件支持把 Kafka的数据直接落到ArgoDB?	144
8.12.8.	Workflow任务打不开	144
8.12.9.	Workflow内置时间参数有哪些?	145
8.12.10.	Sophon添加了quark数据集连接之后, 会显示无状态请更新	145
8.12.11.	StellarDB v4.1版本, 有详细的python Sdk 说明demo么?	145
8.12.12.	执行Spark任务, executor节点报错显示UnknownHost.	146
8.12.13.	TDH上有 kafka的管理界面吗?	146
客户服务		147

免责声明

本说明书依据现有信息制作, 其内容如有更改, 恕不另行通知。星环信息科技(上海)股份有限公司在编写该说明书的时候已尽最大努力保证其内容准确可靠, 但星环信息科技(上海)股份有限公司不对本说明书中的遗漏、不准确或印刷错误导致的损失和损害承担责任。具体产品使用请以实际使用为准。

注释: Hadoop® 和 SPARK® 是ApacheTM 软件基金会在美国和其他国家的商标或注册的商标。Java® 是Oracle公司在美国和其他国家的商标或注册的商标。Intel® 和Xeon® 是英特尔公司在美国、中国和其他国家的商标或注册的商标。

版权所有 © 2013年-2024年星环信息科技(上海)股份有限公司。保留所有权利。

©星环信息科技(上海)股份有限公司版权所有, 并保留对本说明书及本声明的最终解释权和修改权。本说明书的版权归星环信息科技(上海)股份有限公司所有。未得到星环信息科技(上海)股份有限公司的书面许可, 任何人不得以任何方式或形式对本说明书内的任何部分进行复制、摘录、备份、修改、传播、翻译成其他语言或将其全部或部分用于商业用途。

手册版本信息

版本号: T00993x-92-010

发布日期: 2024-6

1. 社区版家族介绍

1.1. 关于社区版您可能想要知道的

a. 什么是社区版？

依托于TDH强大的技术底座，星环科技推出TDH社区版及社区开发版产品，充分降低了用户接触使用大数据技术的使用门槛以及使用成本，用户可以轻松应对海量多源异构数据的高效存储，关联分析等业务需求。

同时，企业用户可以通过订阅的方式，低成本地搭建批处理数据仓库，构建数据湖，从而更深入地洞察业务创新机遇，助力业务决策支持。

b. 社区版有什么优势？

TDH拥有12项领先的基础核心技术，应第三方机构测评，是12年来全球首个通过数据库领域难度最高的基准测试TPC-DS的产品，性能优于cloudera数十倍，优于RDB数百倍。

TDH社区版承载了TDH全部领先的核心技术优势，比如一体多模、统一架构等众多优势，相较于开源方案，拥有“三低一高”等特点（架构复杂度低、开发成本低、运维成本低、数据处理效率高）。

并且，社区版产品在功能上未做裁剪，免费提供了业内极具竞争优势的多个产品。给予了数据全生命周期的安全防护能力及一站式的综合运维管理能力，使用者除了可以借助可视化监控平台对系统负载，平台运行状况等指标进行统一管理与监控，亦可通过平台预置的告警通知等功能实现事前预警、事中告警、事后分析的全阶段运维效果。

为了方便用户使用及安装，社区版的安装流程及部署时间都有大幅度缩减，社区开发版更是将安装步骤降低至三步即可30s一键启动。相较于其他第三方平台，社区版及社区开发版80%的操作均由系统自动化运行，用户仅需跟随系统向导执行即可。

c. 社区版以及社区开发版的区别是什么？

社区版为用户免费提供了4节点10TB的容量支持，用户可以基于最低3台服务器，每台最低4核8GB即可部署专属于用户自己的分布式集群。

为了进一步满足更多用户在资源成本方面的需求，社区开发版通过all-in-one的设计思路，实现了开箱即用的产品特点，针对安装时间、成本均做了大幅度的优化。最低仅需单台服务器，4核16GB/24GB即可。

d. 如果到达容量上限或者想要使用更多组件怎么办？

当前，TDH社区版免费提供了Inceptor关系型分析引擎以及Hyperbase Nosql宽表数据库，为用户解决多样化的数据存储与处理的问题。

如果用户需要更多容量支持以及业务存在其他场景需求，可考虑升级社区订阅版。社区订阅版除了前述的核心组件之外，还为用户提供Scope全文检索、StellarDB图数据库、TDS核心开发工具套件的订阅服务，订阅后即可低成本享受**商业版**组件的核心功能与特性。并且，升级后亦可享受全套的专业服务，如，5*10原厂技术咨询（远程和驻场），集群规划安装扩容，数据平台建设运营，专家诊断优化等等。

如何申请订阅？订阅申请流程：<https://community.transwarp.cn/article/detail?id=329>

家族产品	核心特点	支持的核心组件
 社区开发版	<ul style="list-style-type: none"> ✓ 单机即可部署 	<ul style="list-style-type: none"> Inceptor 关系型分析引擎 Hyperbase NoSQL 宽表数据库 StellarDB 图数据库
	<ul style="list-style-type: none"> ✓ 开箱即用，一键启动 	<ul style="list-style-type: none"> Scope 搜索引擎 Waterdrop 数据库开发工具 星环大数据运维管理套件及 Hadoop 精选组件 ..
 社区版	<ul style="list-style-type: none"> ✓ 资源成本低 	<ul style="list-style-type: none"> Inceptor 关系型分析引擎
	<ul style="list-style-type: none"> ✓ 免费提供 4 节点 10TB 容量支持 	<ul style="list-style-type: none"> Hyperbase NoSQL 宽表数据库
	<ul style="list-style-type: none"> ✓ 分布式集群服务保障服务高可用 	<ul style="list-style-type: none"> Waterdrop 数据库开发工具
	<ul style="list-style-type: none"> ✓ 极致性能与可扩展性 	<ul style="list-style-type: none"> 星环大数据运维管理套件及 Hadoop 精选组件 ..
	<ul style="list-style-type: none"> ✓ 提供轻量级、业务无感知的在线维护 	
 社区订阅版	<ul style="list-style-type: none"> ✓ 无节点容量限制 	<ul style="list-style-type: none"> Inceptor 关系型分析引擎
	<ul style="list-style-type: none"> ✓ 企业级产品、功能及性能 	<ul style="list-style-type: none"> Hyperbase NoSQL 宽表数据库
	<ul style="list-style-type: none"> ✓ 业务规模增长支持 	<ul style="list-style-type: none"> StellarDB 图数据库
	<ul style="list-style-type: none"> ✓ 丰富的业务场景支持 	<ul style="list-style-type: none"> Scope 搜索引擎
	<ul style="list-style-type: none"> ✓ 丰富的业务场景支持 	<ul style="list-style-type: none"> Waterdrop 数据库开发工具
	<ul style="list-style-type: none"> ✓ 专家级集群规划、诊断优化 	<ul style="list-style-type: none"> Transporter 大数据整合工具
	<ul style="list-style-type: none"> ✓ 高级别技术咨询，原厂售后支持 	<ul style="list-style-type: none"> Workflow 任务调度软件 SQLBook (按需提供)



社区版安装手册及教程：<https://transwarp-ce-1253207870.cos.ap-shanghai.myqcloud.com/TDH-CE-2024-5/TDH-CE-InstallManual.pdf>

社区开发版安装手册及教程：<https://transwarp-ce-1253207870.cos.ap-shanghai.myqcloud.com/TDH-CE-2024-5/TDH-CE-DEV-InstallManual.pdf>

2. 安装及使用前须知（重要）

2.1. 安装及使用注意事项

2.1.1. 社区版



创建一个集群前，首先必须对整个集群进行规划，规划包括集群节点、网络、软件模块分布等部分。

请社区版用户在安装前再次进行以下检查，以确保未来在使用产品的过程中可以减少大量的风险和阻碍。

1. 硬件检查

jdk1.8

CentOs 7.3-7.9

Red Hat Enterprise Linux 7.3-7.9

必须是3台以上服务器组成，每台服务器最低配置如下：

- 4核心或以上带超线程x86指令集CPU的服务器
- 8GB以上内存
- 2个300G以上的硬盘做RAID1，作为系统盘
- 4个以上的300GB容量以上的硬盘作为数据存放硬盘
- 2个千兆以上网卡



注意事项： 1) 上述配置仅可满足基础组件的安装以及进行基础功能测试，不支持跑批量数据，测试进阶功能等场景，如需体验更多功能或使用更多组件，如Hyperbase/Guardian等，为实现更好的性能，请基于自身需求将配置增加至至少4核/8核32-64GB及以上； 2) 如果只是验证功能，单做测试搭建，而且数据量不大的话docker盘和数据盘也可适当将降低至各100G左右；

2. 冲突包检查（务必保证）

是否存在与 `tos, manager` 安装相冲突的包

冲突包list:

- 检查manager所在节点是否已有mysql，如果有建议卸载，如果不想卸载，可更换一个干净的节点安装Manager；
- 检查原操作系统是否有自带的docker，如果有，建议卸载再进行安装；
- 检查container-selinux是否冲突，如果有冲突建议执行`rpm -qa | grep selinux`然后执行，`rpm -e container-selinux`，可能会提示有依赖，需要逐步把依赖的包都卸载掉；

3. docker 分区挂载检查（务必保证）

须确保单独挂载 `/var/lib/docker` 目录

4. docker 分区大小检查（务必保证）

/var/lib/docker 分区大于100G

5. docker 分区检查（务必保证）

/var/lib/docker 分区 ftype = 1

6. sudo 用户检查（务必保证）

安装用户需要具有 sudo 权限，并且NOPASSWD，或是root用户

7. 根分区大小检查（推荐）

推荐大于100GB

8. 日志分区挂载检查（推荐）

单独挂载 /var/log 目录

9. 日志分区大小检查（推荐）

/var/log 分区推荐大于100G

10. Manager浏览器支持说明：

Google Chrome（推荐）：71及以上

FireFox：64及以上

Safari：11.1.2及以上，或12.0.2及以上

Edge：79及以上

11. 一键开启安全，杜绝挨个开启

Guardian开启安全时需要一键开启，不建议一个一个的开启组件安全。同时，TDH 有统一的认证管理不要再用开源查到的参数配置Krb5。

此外需要注意，开启安全或关闭安全均需要重新下载TDH-Client才可以包含开启安全或关闭安全后的配置。



12. 出现大面积宕机，优先解决核心基础服务

TOS/License/guardian/transwarp-manager-agent，解决后可以看每个组件的依赖关系图。底层的基础服务正常了，就可以重启这个服务了。

通常大致顺序是：zk→hdfs→yarn→hyperbase→quark，具体还需查看自己的集群服务情况。

13. 操作雷区，避免此类操作

删除核心文件，比如transwarp-manager-common等；

删除核心服务，比如manager中的mysql等；

14. 检查下载的产品包文件是否被损坏

与产品包中带的md5文件或sha256文件进行对比，文件里包括一个md5或sha256值，用于查看产品包是

否被损坏。

md5文件校验命令指令: md5sum <pkg>
sha256文件校验命令指令: sha256sum <pkg>

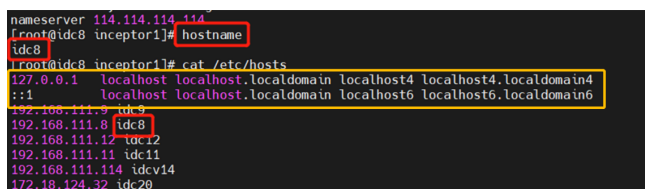
例.



一致，说明未被损坏

15. 登入到节点使用 hostname 命令检查节点主机名

确保节点名称与上述 /etc/hosts 里设置的名称相同，如下图所示：



16. 安装包说明

您从官网下载的产品包进行解压后不同产品包对应的名称：

Manager安装包： TDH-Platform-Community-Transwarp-2024.5-X86_64-final.tar.gz

基础产品包： TDH-CE-ImageFiles-transwarp-9.3.3-ce-final.tar.gz

基础产品包中仅包含Inceptor/AquilaInsight/Hadoop/KunDB/TOS服务，如需Hyperbase/Guardian组件请从官网下载弹窗下载



请在进行安装过程前提前将Manager产品包以及子产品安装包上传至服务器

2.1.2. 社区开发版



- 1) 同一台机器上不可同时安装社区版以及社区开发版，同时运行可能会出现端口冲突的情况；
- 2) 社区开发版在初次安装时已经安装好了全部的服务，因此不支持在线上传产品包或添加服务；

当前社区开发版系列产品分为：

- TDH社区开发版
- StellarDB社区开发版
- Scope社区开发版

上述三个开发版产品当前版本暂不支持统一平台统一纳管，预计将于下个版本做更新，因此当前如果需要安装启动多个开发版产品需要仔细查看下方操作事项。

1. 本地目录的挂载点需不同（务必确保）

为了实现数据的共享以及持久化，社区开发版系列产品当前需要把自定义的一个本地目录挂载到容器内的/opt/transwarp目录下。挂载后服务的配置，日志，数据以及服务的数据目录都会放在这个本地目录下。

因此会有一个统一格式的命令用于启动容器及挂载相关目录：

```
docker run -d --network host -v <本地目录路径>:/opt/transwarp --privileged <镜像名称>
```

为了防止多个产品的相关数据挂载到同一目录导致数据串掉或丢失，启动所用的<本地目录路径>不同产品需要分别采用不同的目录，例如：

```
/home/dev_tdh
```

```
/home/dev_stellardb
```

```
/home/dev_scope
```



请注意，在安装新的产品包时，请确保挂载的本地目录下是否已存在社区开发版相关的数据文件，请务必选择空目录进行挂载

升级除外，升级请看升级手册：<https://transwarp-ce-1253207870.cos.ap-shanghai.myqcloud.com/TDH-CE-2024-5/TDH-CE-DEV-UpgradeManual.pdf>)

2. 如果您需要运行多个开发版容器，在使用某一开发版前，请先停止其他开发版的容器30s后，再启动新的容器，以防止端口冲突。

相关命令：
`docker ps`
`docker stop <other_containerid>`
`docker start <new_containerid>`

3. 在使用完某一开发版产品，更换产品使用时，如果8180页面未加载出来对应产品的manager平台，可尝试清除缓存或使用ctrl+F5强制刷新页面。

4. 请勿随意更改容器内的权限，如果需要为用户赋权，下载客户端执行初始化后可基于下述两种情况在容器外赋权：

a. 如果开启guardian，在guardian上做赋权（给hive用户赋予/目录权限）

b. 未开启guardian（建议开启guardian），通过下方命令来做赋权：

```
export HADOOP_USER_NAME=hdfs
hdfs dfs -chmod -R 777 /
```

客户端下载及使用请参考安装手册中其他使用方面的操作章节

5. 如果在安装过程中不小心关闭页面或因为其他原因导致退出进程，可重新访问Manager页面进入运维诊断页面选择具体的进程继续即可。

集群管理 运维诊断 应用市场 系统管理

搜索... Aquila 全局服务

操作

查询和查看提交执行的操作。

* 操作时间: 2024-06-13 15:12:23 → 2024-06-13 17:00:00 资源: 请输入 用户名: 请输入

状态: 请选择

查询 收起

操作名称	状态	提交时间	开始时间	结束时间	资源	用户名	用户IP
初始化KGExplorer1	SUCCESSFUL	2024-06-13 15:35:29	2024-06-13 15:42:54	2024-06-13 15:43:59	KGExplorer1	admin	172.16.201.84
初始化Quark1	SUCCESSFUL	2024-06-13 15:35:29	2024-06-13 15:41:14	2024-06-13 15:42:54	Quark1	admin	172.16.201.84
初始化YARN1	SUCCESSFUL	2024-06-13 15:35:29	2024-06-13 15:40:03	2024-06-13 15:41:14	YARN1	admin	172.16.201.84
初始化StellardbStorage1	SUCCESSFUL	2024-06-13 15:35:29	2024-06-13 15:39:48	2024-06-13 15:40:03	StellardbStorage1	admin	172.16.201.84
初始化HDFS1	SUCCESSFUL	2024-06-13 15:35:29	2024-06-13 15:39:13	2024-06-13 15:39:48	HDFS1	admin	172.16.201.84
初始化ZooKeeper1	SUCCESSFUL	2024-06-13 15:35:28	2024-06-13 15:39:00	2024-06-13 15:39:13	ZooKeeper1	admin	172.16.201.84
初始化Guardian	SUCCESSFUL	2024-06-13 15:35:28	2024-06-13 15:37:23	2024-06-13 15:39:00	Guardian	admin	172.16.201.84
初始化Aquila	SUCCESSFUL	2024-06-13 15:35:28	2024-06-13 15:35:43	2024-06-13 15:37:23	Aquila	admin	172.16.201.84

3. 常见运维指令

3.1. 社区版Manager操作指令集

如何查看状态、启停

```
Manager支持以下几种指令： status | stop | start | restart
```

a. Agent进程

```
systemctl <status | stop | start | restart> transwarp-manager-agent
```

b. master进程

```
systemctl <status | stop | start | restart> transwarp-manager
```

c. Manager自身使用的mariadb

```
systemctl <status | stop | start | restart> transwarp-manager-db
```

3.2. 社区版开发版镜像&容器相关指令

- docker images 查镜像
- docker ps -a 查容器（容器是基于镜像创建）
- docker exec -it [容器id] bash 进入容器内
- docker start 启动容器

更详细的内容可查看后续<常见的Docker指令>小节

3.3. 常见的Linux运维指令

3.3.1. 目录操作

- bin: 存放最常用的命令
- boot: 存放linux启动的核心文件
- home: 用户的主目录
- dev: 设备的缩写，存放的是linux的外部设备
- lost+found: 存放非法关机文件
- etc: 存放系统管理所需要的配置文件
- usr: 重要软件存放地点相当于windows 的program file
 - src: 内核源代码默认的放置目录
 - bin: 用户使用的应用程序

- opt: 主机额外安装的摆放的目录, xmind、redis、Oracle等
- var: 扩张的内存存放
- run: 存放临时文件, 当系统重启时会被删除
- temp: 临时文件存放目录用完即丢
- www: 存放服务器网站的相关资源

3.3.2. 文件系统以及权限

ll显示文件属性和权限, linux是一个多用户系统, 所以权限繁多

- [d] : 目录
- [-] : 文件
- [l] : 链接文档
- [b] : 表示为装置文件里面可供存储的接口设备
- [c] : 表示为装置文件的串行设备。例如: 鼠标、键盘等
- r: 代表可读 read
- w: 代表可写 write
- x: 代表可执行 execute
- 0位: 文件类型 d、-、l、b、c
- 1-3位: 属主权限 ——root可读、可写、可操作
- 4-6位: 属组权限——其他用户无法写
- 7-9位: 其他用户权限——其他用户无法写

3.3.2.1. 更改文件属组

```
chgrp [-R] 属组名 文件名
```

- 参数 r 是递归修改

3.3.2.2. 更改文件属主

```
chown [-R] 属主名 文件名
```

3.3.2.3. 更改文件的9个属性

```
chmod[-R] 文件或目录
```

Linux文件属性有两种设置方法

- 符号: -rwxrwxrwx
- 数字

r:4

w:2

x:1

可读可写不可执行 `rw-` 6

可读可写可执行 `rwx` 7

不可读不可写不可执行 `---` 0

每种身份各自对应三个权限

`owner = rwx = 4+2+1 = 7`

`group = rwx = 4+2+1 = 7`

`others = rwx = 4+2+1 = 7`

`chmod 777 filename`

3.3.3. 文件和目录操作

- `ls`: 列出目录内容。
- `cd`: 切换目录。
 - 绝对路径 `/`
 - 相对路径 `../`
 - 用户目录 `~`
- `pwd`: 显示当前工作目录。
- `mkdir`: 创建一个目录。
 - `-p`: 循环创建目录
- `rmdir`: 删除目录。
 - `-p`: 层级删除内部文件
- `rm`: 移除文件或者目录
 - `-f`: 忽略不存在的文件，不会出现警告，强制删除
 - `-r`: 递归删除目录
 - `-i`: 互动，询问是否删除
- `cp`: 复制文件或目录。
- `mv`: 移动文件或目录。
- `touch`: 创建空文件或更新文件的访问时间戳。

3.3.4. 文本处理，查看文件内容

- `cat`: 由第一行开始显示文件内容。
- `tac`: 从最后一行开始显示，可以看出`tac`是`cat`的倒写
- `grep`: 在文件中搜索指定模式。
- `sed`: 流编辑器，用于文本替换和处理。
- `awk`: 文本处理工具，支持文本提取、格式化等操作。
- `nl`: 显示的时候，可以输出行号
- `more`: 一页一页地显示文件的内容
 - 空格是翻页
 - `enter`是向下翻一行

退出是q命令

- less: 与more类似但是可以往前面翻页
- head: 只看头几行
- tail: 只看尾部几行



可以使用man 命令来查看各个命令的使用文档

3.3.5. 进程管理

- ps: 显示当前进程状态。
- top: 实时显示系统资源使用情况和进程信息。
- kill: 终止指定进程。
- pkill: 根据进程名终止进程。

3.3.6. 用户和权限管理:

- useradd: 创建新用户。
- passwd: 设置用户密码。
- usermod: 修改用户属性。
- userdel: 删除用户。
- chown: 修改文件或目录的所有者。
- chmod: 修改文件或目录的权限。

3.3.7. 系统信息

- uname: 显示系统信息。
- hostname: 显示主机名。
- df: 显示磁盘空间使用情况。
- free: 显示内存使用情况。
- uptime: 显示系统运行时间和平均负载。

3.3.8. 网络管理

- ifconfig 或 ip addr: 显示网络接口信息。
- ping: 测试网络连接。
- netstat: 显示网络状态和连接信息。
- traceroute: 跟踪数据包的路径。

3.3.9. 包管理

- apt (Debian/Ubuntu): 用于包管理、安装、更新和删除软件包。
- yum (CentOS/RHEL): 用于包管理、安装、更新和删除软件包。

3.3.10. Linux的链接

- 硬链接：A---B，假设B是A的硬链接。那么他们两个指向了同一文件！允许一个文件拥有多个路径，一般用于重要文件上，防止误删。
- 软链接：类似于Windows下的快捷方式，源文件一旦删除，软链接将无法访问

创建硬链接

```
ln file1 file2 创建硬链接
```

创建软链接

```
ln -sf sourcefile softsourcefile
```

3.3.11. Vim编辑器基本操作

- vim基本上分为三种模式

命令模式

输入模式

home/end : 移动光标到行首行位

PageUp/PageDown : 上下翻页

底线命令模式

n :跳转到第n行:

/word : 向光标之下寻找word

? word: 向光标之上寻找word

n/N:基础上寻找下一位

w: 写入

q: 退出

!: 强制

wq: 先写入，后离开

ZZ: 如果变更写入离开，如果没变更直接离开

set nu: 显示行号

set nonu : 取消行号

3.3.12. 账号管理

linux是一个多用户的分时管理系统，用户账号的管理工作主要涉及到用户账号的添加、删除、修改

- 添加用户

```
useradd -选项 用户名
```

-m: 自动创建这个用户的主目录/home

-G: 分配一个用户组

- 删除用户

userdel -r 用户名

- 修改用户

usermod -参数 用户名

-d: 用户主目录

-g: 用户组目录

- 切换用户

su username

- 锁定用户

passwd -l username

passwd -d username 将密码清空，无法下次登录

以上是一些常用的 Linux 运维指令，涵盖了日常系统管理和操作的基本需求，如需了解更多详细指令，可前往星环开发者社区了解：<https://community.transwarp.cn/article/detail?id=399>

3.4. 常见的Docker指令

3.4.1. 常用的基础操作

- 容器状态介绍

created: 已创建

restarting: 重启中

running: 运行中

removing: 迁移中

paused: 暂停

exited: 停止

dead: 死亡

- 容器管理

docker ps: 列出正在运行的容器。

docker ps -a: 列出所有状态的容器，包括正在运行的和已停止的。

docker stop <container_id>: 停止运行容器。

docker restart <container_id>: 重启容器。

docker rm <container_id>: 删除已停止的容器。

docker logs <container_id>: 查看容器的日志，-f参数实时刷新日志。

docker logs -f --tail=500 <container_id>: 实时查看最后500条日志；

docker logs --since 30m <container_id>: 查看最近30分钟的日志；

docker exec -it <container_id> <command>: 在运行中的容器中执行命令。

docker exec -it XXX sh: 使用sh命令进入运行中的容器的控制台中



- a. 容器名不可重复，不会覆盖，创建新版本的同名容器之前需要先停止并删除旧容器
- b. 可以使用容器id或容器名作为容器的唯一标识，命令中可以更换使用

- 运行:

1. 启动container, 运行某一个image: `docker run -it <image_id or image_tag> bash`
2. 提交image的修改: `docker commit <container_id> <image_tag>`
3. push image的修改: `docker push <image_tag>`

3.4.2. 其他进阶操作

- 容器使用

```
$ docker run -it ubuntu /bin/bash
```

-i: 交互式操作

-t: 终端

ubuntu: 镜像名

/bin/bash: 启动 shell 交互式命令

-d: 后台运行

- 镜像管理

`docker images`: 列出本地的存在的镜像列表。

`image_tag` 若为空, 则使用默认的latest tag: `docker pull image_name:image_tag`

注: 在本地的docker操作时, 镜像id和镜像名+tag均可作为容器的唯一标识, 在命令中可替换使用

`docker pull <image_name>`: 从仓库拉取镜像。

`docker tag XXX new_tag`: 为镜像打标签, 若标签已存在, 会将标签赋给新镜像, 原镜像标签为空 (悬空镜像)

镜像输出为tar包:

```
docker save XXX -o name.tar
```

```
docker save XXX > name.tar
```

导入镜像包:

```
docker load -i name.tar
```

```
cat name.tar > docker load
```

`docker rmi <image_id>`: 删除镜像。

注: 在docker build 的过程中会出现中间镜像。中间镜像不一定会自动删除。会占用存储空间

可以通过`docker images -f dangling=true`命令查找悬空的镜像。

```
docker rmi $(docker images -f "dangling=true" -q): 删除悬空镜像。
```

```
docker build <path_to_dockerfile>: 根据 Dockerfile 构建镜像。
```

- 网络管理

`docker network ls`: 列出 Docker 网络。

`docker network inspect <network_id>`: 查看特定网络的详细信息。

- 数据卷管理

`docker volume ls`: 列出 Docker 数据卷。

`docker volume create <volume_name>`: 创建数据卷。

`docker volume inspect <volume_name>`: 查看数据卷的详细信息。

- 系统管理:

`docker info`: 显示 Docker 系统信息。

`docker version`: 显示 Docker 版本信息。

`docker system df`: 显示 Docker 使用的磁盘空间信息。

`docker system prune`: 清理未使用的资源（容器、镜像、数据卷等）。

以上是一些常用的Docker指令，你可以根据具体的需求进一步了解和使用的。

3.5. 常见的K8S命令

本篇章主要讲述在处理pod故障时的常用命令。

- 查看某个资源

```
kubectl get {pod|node|deploy|rs} {资源名} [-o [wide|yaml|json]] [-n default|kube-system]
[资源名选填，如果写上代表只获得这个资源的具体信息]
```

`-o wide` 代表扩展pod信息，增加了 NODE 和 IP 信息

`-o yaml|json` 代表以 `yaml|json` 形式展示 pod 的信息

`-n` 选填，不选表示所有 default namespace 下的资源

例如：查看 default namespace 下的所有 hdfs 相关的 pod 资源

```
[root@tdh-01~]# kubectl get pods -o wide |grep hdfs
```

- 查看默认namespace下的pod

```
kubectl get pods
```

- 查看某个pod的pod日志

```
kubectl logs {pod名} [-f]
```

`-f`: 表示实时动态展示具体pod的pod日志

```
# 例如：查看 podname 为 inceptor-server-inceptor1-3529736423-qh58n 的 pod 日志
[root@tdh-01~]# kubectl logs inceptor-server-inceptor1-3529736423-qh58n
```

- 查看某个资源的详细信息

```
kubectl describe {pod|node|deploy|rs} {资源名}
```

该命令可以获取某个资源的具体信息，该资源可以是pod, node, deployment, replicaset等

例如：查看 podname 为 inceptor-server-inceptor1-3529736423-qh58n 的 pod 描述信息以及事件信息

```
[root@tdh-01~]# kubectl describe pods inceptor-server-inceptor1-3529736423-qh58n
```

- 删除某个资源

```
kubectl delete {pod|node|deply|rs} {资源名}
```

例如：删除 podname 为 inceptor-server-inceptor1-3529736423-qh58n 的 pod

```
[root@tdh-01~]# kubectl delete pods inceptor-server-inceptor1-3529736423-qh58n
```

- 删除所有资源（慎用慎用慎用哦）

```
kubectl delete {pod|node|deply|rs} --all
```

- 删除所有非Running的pod（慎用慎用慎用哦）

```
for i in `kubectl get pod -o wide | grep -v Running | grep -v NAME | awk '{print$1}'`; do  
kubectl delete pod ${i}; done
```

- 编辑某个资源（不熟悉的话不建议用哦）

```
kubectl edit {pod|node|deploy|rs} {资源名}
```

- 运行

1. 运行默认namespace下的pod:

```
kubectl exec -it <pod_id> bash
```

2. 运行指定namespace下的pod

```
kubectl exec -it <pod_id> bash --namespace <namespace_name>
```

4. 安装前可能想要了解的产品问题

4.1. TDH社区开发版与TDH社区版能否部署在同一服务器?

【用户问题】

可不可以一台服务器既装社区开发版也装社区版，一个用来测试，一个用来日常使用?

【问题答复】

不可以的，因为可能会产生端口冲突影响后续的产品使用，如果一定一定要装到一起的话需要手动停很多服务比如transwarp-manager-agent、transwarp-manager-master等等，所以建议不要混着装，最好用不同的机器分开安装。

4.2. 星环科技是否提供 Hive ?

【用户问题】

星环科技是否提供 Hive ?

【问题答复】

星环Inceptor不运行在Hive的代码之上，只是因为需要对接Hadoop的生态系统，所以在接口部分包含了部分Hive的源码，对不同版本Hive的语法做了大部分的兼容。

4.3. Quark是什么?

【用户问题】

Quark是什么?

【问题答复】

Quark是星环自研的SQL编译器，可以确保星环平台中的各个数据库产品遵从一致的SQL规范。

4.4. ArgoDB/Inceptor/Hyperbase三者有什么区别?

【用户问题1】

ArgoDB/Inceptor/Hyperbase有什么区别?

【问题答复】

- 先说Hyperbase，Hyperbase主要采用KV存储格式，因此更适合精确查询、点查的操作场景，不适合应用于批量数据的加工处理。
- ArgoDB与Inceptor均可以应用于数据批量处理等场景，都可以满足用户批处理、数据仓库、湖仓一体等业务需求。而且在最新的TDH9产品里，Inceptor的最高级版本可以实现湖仓集一体的业务场景。

不过，ArgoDB在计算性能以及引擎上有进一步的升级，是Inceptor的Plus。ArgoDB采用了基于C++的OLAP引擎，基于湖仓基本能力做了进一步升级，最高版本除了可以实现包含Inceptor能处理的所有场景之外还多了实时数仓，数据集市，历史数据查询等场景的能力，用户可以低延时的处理数据的插入更新，也可以满足用户高性能并发在线OLAP需求。

【用户问题2】

argodb函数、sql语法和inceptor一样吗？

【问题答复】

- 在函数和SQL语法方面，Argodb函数和Inceptor绝大部分都是一样的，DDL略有不同，Query语法一样。

【用户问题3】

inceptor对交易型业务的支持性能怎么样？比如数据查询之类的

【问题答复】

- 在对交易型业务的支持方面。Inceptor不支持交易型业务，该组件主要擅长做批处理和数据加工，数据查询的一般是Hyperbase或者ArgoDB。Hyperbase是适合精确查询并且返回结果不多，ArgoDB则支持面更大一些。

【用户问题4】

如果离线处理的较多，不涉及实时计算，基础的Inceptor可以基本满足需求吗？

【问题答复】

- ArgoDB主要依赖Holodesk作为存储格式，和Inceptor公用了Quark，但是某些算法是只有Holodesk上才支持，如果只跑批的话，Inceptor是可以满足需求的。

4.5. TDH集群可以部署两个HDFS服务和Yarn服务吗？

【用户问题】

请问在一个Manager中，可以部署两个HDFS和Yarn吗？Yarn1只访问HDFS1，Yarn2只访问HDFS2。

【问题答复】

Manager是支持部署多个服务的，该部署方式需配置依赖关系，Yarn1依赖HDFS1，Yarn2依赖HDFS2。

4.6. Manager可以添加多个Quark服务吗？

【用户问题】

Manager可以添加多个Quark服务吗？

【问题答复】

Quark是可以按需配置多个的，不仅可以用于高可用的场景，也可以基于集群规划进行资源分配，按照业务切等等都可以。

4.7. TOS Registry可以多活部署吗？

【用户问题】

请问TOS Registry可以多活部署吗？印象中好像只能部署一个，在Manager节点？

【问题答复】

您好，是支持独立部署多个Registry来避免单点故障风险的。具体的RegistryHA方案可以咨询星环专业技术团队。

4.8. 社区开发版想换成社区版，使用同一节点需要什么操作吗？

【用户问题】

社区开发版想换成社区版，使用同一节点需要什么操作吗？需要删除docker吗？

【问题答复】

容器停掉，挂载目录删掉就可以，Docker想继续用可以不删。

4.9. 如果数据源都是通过API接口方式获取，星环平台有什么工具是可以用作数据采集的吗？

【用户问题】

如果数据源都是以API接口调用的方式获取，请问可以用哪种方式将数据采集到HDFS。目前待验证的思路是开发代码将数据采集到kafka，然后再将数据消费到Inceptor。请问还有其他更加高效的方式吗，比如说功能组件等直接可以支持API作为数据源？

【问题答复】

您好，目前往HDFS写数据，最多的方式就是使用API的方式，这里的API是指hdfs提供的client jar包里的提供的。您提到的将数据采集到kafka，然后再将数据消费到inceptor是否可行主要取决于，您是想存到inceptor还是hdfs，特别是以什么样的形式来保存。

其次我们有一个TDS工具是Transporter（大数据整合工具），可以支持将分散于各个地方、各种平台上的各种格式的数据同步或集成到大数据平台上，通过简洁、统一的可视化界面快速配置数据流转流程，实现异构平台和数据源之间的数据流转。

目前这个产品已在我们的社区订阅版中提供，具体信息可查看：<https://community.transwarp.cn/article/1091>

4.10. 像上层组件(Hybase,Scope,StellarDB..)这些最终数据的存储，其实也是落地在Inceptor中，建立相应的数据映射，是否可以这样理解？

【用户问题】

像上层组件(Hybase,Scope,StellarDB..)这些最终数据的存储，其实也是落地在Inceptor中，建立相应的数据映射，是否可以这样理解？

【问题答复】

Inceptor可以访问Hyperbase，以及hadoop生态里的外表，torc等，但是不能访问StellarDB，ArgoDB的存储格式，StellarDB是需要有专用的quark来访问的，不是Inceptor的quark。同理Scope，ArgoDB的quark也是这样的。但是StellarDB的quark是可以访问Inceptor能访问的存储格式，也就是说它可以向下兼容，基础的不可以向上兼容。

Inceptor访问的都是以hdfs为主，而像StellarDB，Scope，ArgoDB，Timelyre包括新发布的产品Hippo这些格式最终的文件都不在hdfs，而是落在星环自己的分布式存储上。这是因为hdfs的问题是只能存数据不能存状态。

顺便，如果想要了解星环分布式数据管理系统的更多信息，可以查看这篇文章：
<https://community.transwarp.cn/article/942>

4.11. 星环可以对外提供数据服务吗，外部程序如何获取数据

【用户问题】

星环大数据平台和外部程序交互的方式是什么，外部程序如何去获取星环大数据平台的数据

【问题答复】

建议使用jdbc连接数据库的方式获取平台数据。或者使用tds工具集中的midgard以restful api的方式发布服务，供外部程序调用；

交互的话 譬如spark，可以将我们相关的配置文件(hive-site, hdfs-site, core-site, yarn-site等)加入spark的conf目录中，然后按照一定要求修改外部应用的部分配置文件即可。

4.12. 社区版对国产化支持度怎么样？可以在麒麟V10里安装吗？

【用户问题】

社区版对国产化支持度怎么样？可以在麒麟V10里安装吗？

【问题答复】

您好，社区版目前暂时不支持，仅支持RedHat以及CentOS操作系统。

但是社区订阅版以及企业版针对一些主流的国产化CPU/操作系统都有支持，比如银河麒麟、统信、openEuler 等等，还有像鲲鹏、飞腾也都有支持。具体可以访问我们的手册以及官网进行查看，如果有一些不支持的您也可以提出您的需求，我们有专门的团队会进行评估进行适配。

5. 【TDH社区版】安装Manager向导过程中可能遇到的问题

5.1. 【RPM配置教程】安装向导显示需要配置RPM库，如何配置？



选择— 使用远程RPM库 — 需要根据自己的操作系统修改变量

```
#
[base]
name=CentOS-$releasever - Base - mirrors.aliyun.com
failovermethod=priority
baseurl=http://mirrors.aliyun.com/centos/$releasever/os/$basearch/
http://mirrors.aliyuncs.com/centos/$releasever/os/$basearch/
http://mirrors.cloud.aliyuncs.com/centos/$releasever/os/$basearch/
gpgcheck=1
gpgkey=http://mirrors.aliyun.com/centos/RPM-GPG-KEY-CentOS-7

#released updates
[updates]
name=CentOS-$releasever - Updates - mirrors.aliyun.com
failovermethod=priority
baseurl=http://mirrors.aliyun.com/centos/$releasever/updates/$basearch/
http://mirrors.aliyuncs.com/centos/$releasever/updates/$basearch/
http://mirrors.cloud.aliyuncs.com/centos/$releasever/updates/$basearch/
gpgcheck=1
gpgkey=http://mirrors.aliyun.com/centos/RPM-GPG-KEY-CentOS-7

#additional packages that may be useful
[extras]
name=CentOS-$releasever - Extras - mirrors.aliyun.com
failovermethod=priority
```

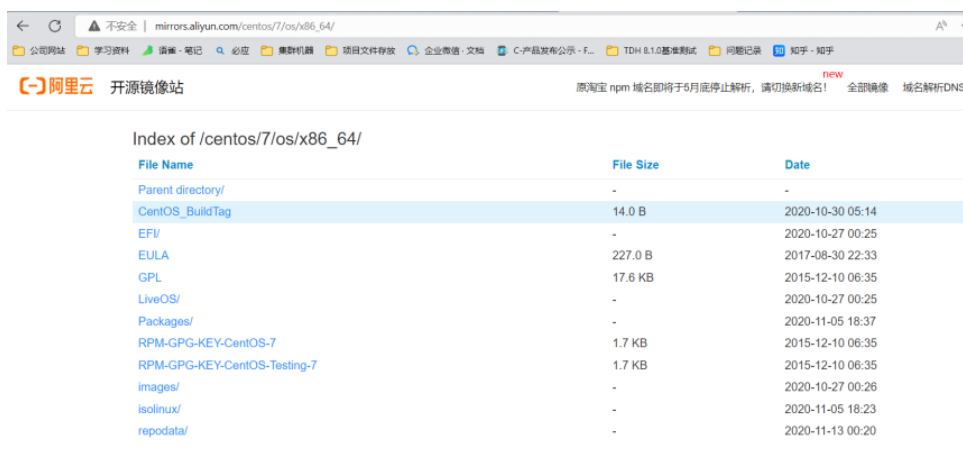
操作示例：

查看自己的操作系统后可以看出：\$releasever的值是：7， \$basearch值是x86_64

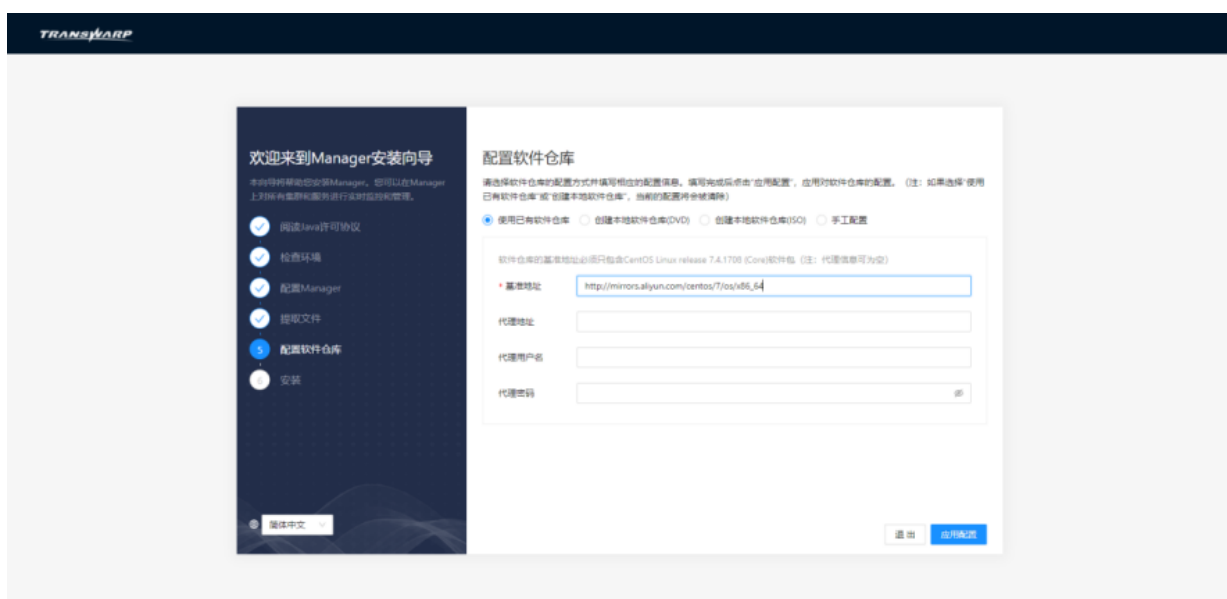

```
[root@intel11 ~]# rpm -qi centos-release
Name      : centos-release
Version   : 7
Release   : 6.1810.2.el7.centos
Architecture: x86_64
Install Date: 2021年08月23日 星期一 20时23分21秒
Group     : System Environment/Base
Size      : 41271
License   : GPLv2
Signature : RSA/SHA256, 2018年11月23日 星期五 21时35分20秒, Key ID 24c6a8a7f4a80eb5
Source RPM : centos-release-7-6.1810.2.el7.centos.src.rpm
Build Date : 2018年11月23日 星期五 21时16分58秒
Build Host : x86-01.bsys.centos.org
Relocations : (not relocatable)
Packager  : CentOS BuildSystem <http://bugs.centos.org>
Vendor    : CentOS
Summary   : CentOS Linux release file
Description :
CentOS Linux release files
[root@intel11 ~]# arch
x86_64
[root@intel11 ~]#
```

基础路径: `baseurl=http://mirrors.aliyun.com/centos/$releasever/os/$basearch/`

那么完整路径则是: `baseurl=http://mirrors.aliyun.com/centos/7/os/x86_64/` 可以访问网页验证是否转换成功



成功后则在安装时输入对应的仓库地址即可

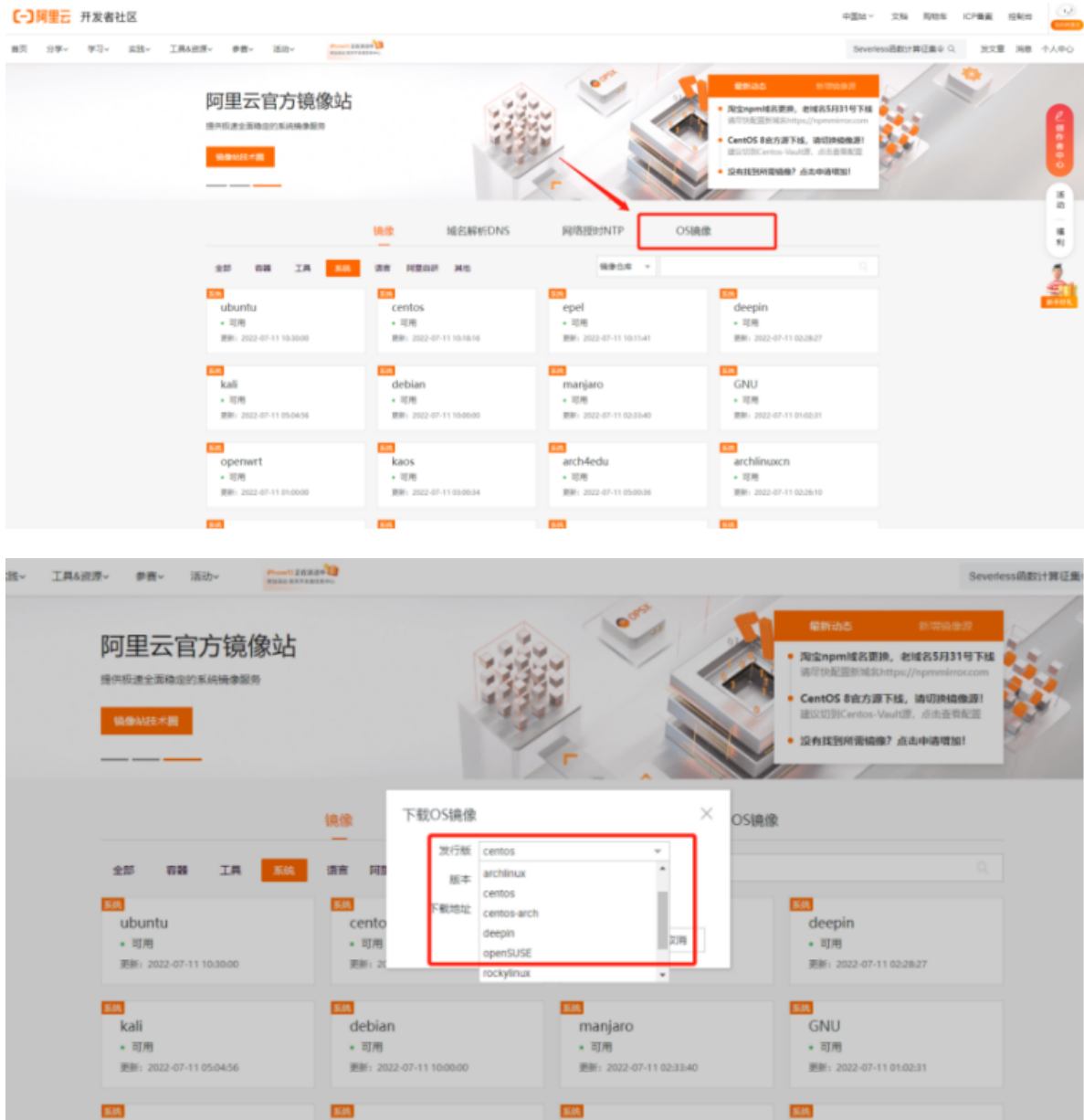


选择二 创建新的RPM库（ISO）【无需联网】

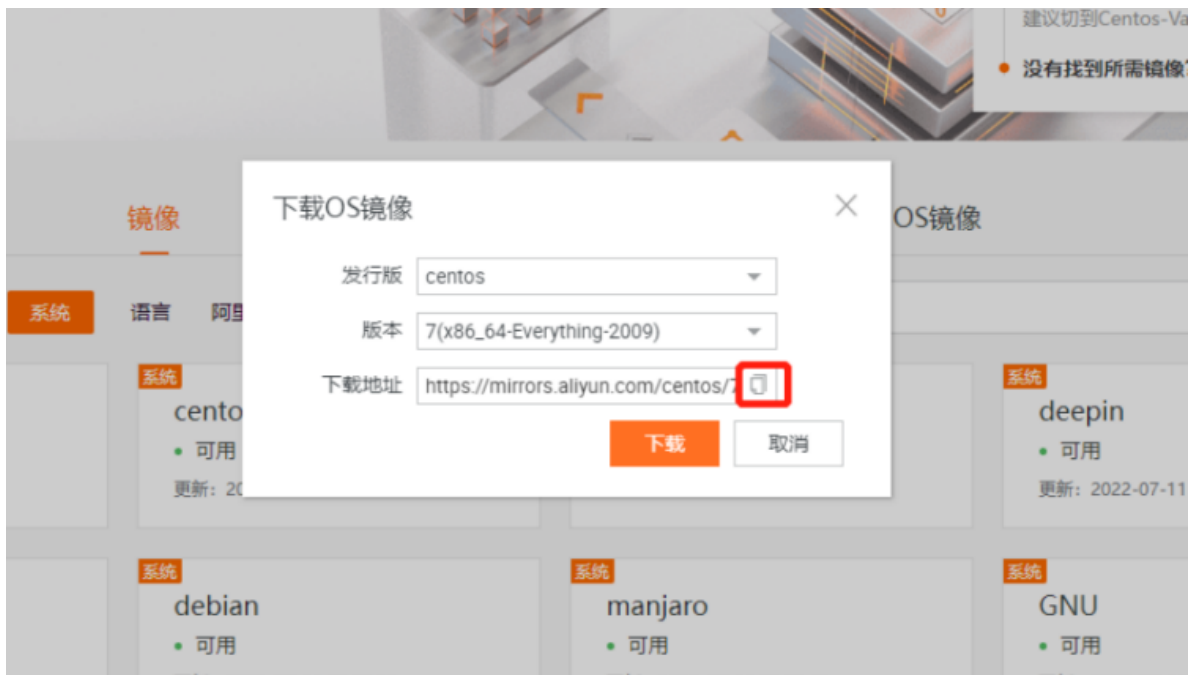
1. 下载ISO镜像文件

下载地址：<https://developer.aliyun.com/mirror/>

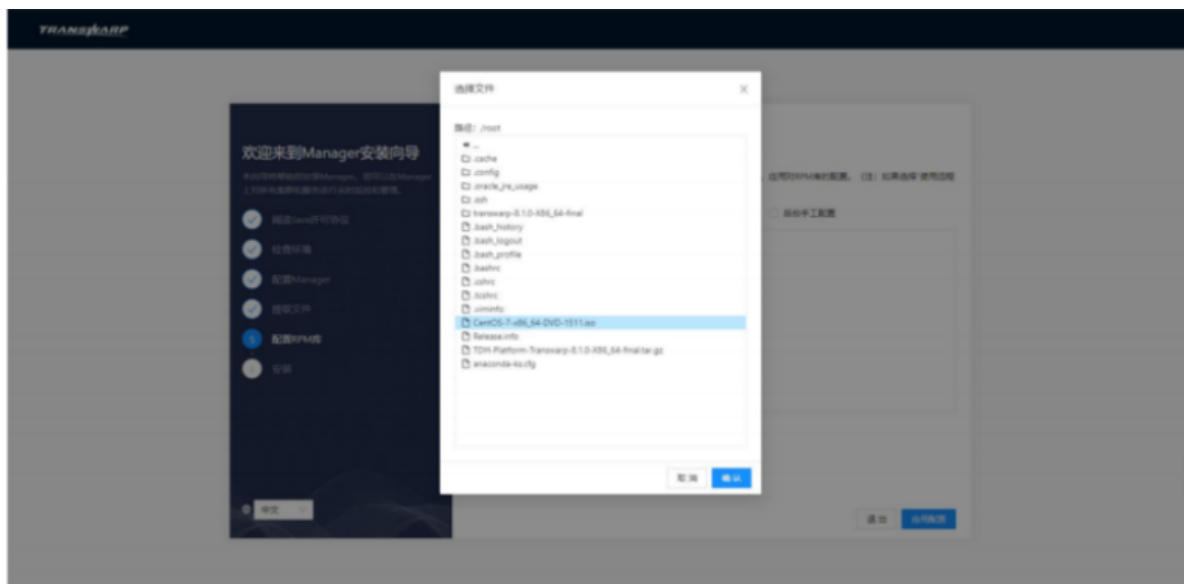
2. 选择对应操作版本的OS镜像（注：社区版要求CentOS版本为7.3-7.9），如：



3. 选择好后复制链接，进行下载



4. 下载好后保存至机器中，安装manager时在该步骤中选择存放在路径下的ISO镜像文件即可



选择三 后台手工配置（CentOS需要能够连接外网）

如果用户没配好yum，但是有一个远程地址可以用，比如阿里云的yum源，就可以选择把地址写到远程仓库的基准地址这里也行

1. 访问官方yum源地址下载对应版本的yum源

地址：<https://developer.aliyun.com/article/675241>

CentOS yum repo

```
# CentOS 5
wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-5.repo
# CentOS 6
wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-6.repo
# CentOS 7
wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-7.repo
```

2. 运行对应版本的命令

例如CentOS7, 运行

wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-7.repo

```

Last login: Mon Jul 11 10:16:29 2022 from 172.16.102.19
*****
# Transwarp PV Cluster
# maintained by DE Tech Platform Division
# - use command "tinit" to initialize TDH-Client
*****
[root@intell1 ~]# wget -O /etc/yum.repos.d/CentOS-Base.repo http://mirrors.aliyun.com/repo/Centos-7.repo
--2022-07-11 10:16:45-- http://mirrors.aliyun.com/repo/Centos-7.repo
正在解析主机 mirrors.aliyun.com (mirrors.aliyun.com)... 118.212.138.238, 118.212.138.248, 118.212.138.217, ...
正在连接 mirrors.aliyun.com (mirrors.aliyun.com)|118.212.138.238|:80... 已连接。
正在发出 HTTP 请求... 正在等待响应... 200  OK
长度: 2523 (2.5K) [application/octet-stream]
正在保存至: "/etc/yum.repos.d/CentOS-Base.repo"

100%[*****] 2,523  --K/s 用时 0.63s

2022-07-11 10:30:45 (82.7 KB/s) - 已保存 "/etc/yum.repos.d/CentOS-Base.repo" [2523/2523]

[root@intell1 ~]# cat /etc/yum.repos.d/CentOS-Base.repo
#
# CentOS-Base.repo
#
# The mirror system uses the connecting IP address of the client and the
# update status of each mirror to pick mirrors that are updated to and
# geographically close to the client.  You should use this for CentOS updates
# unless you are manually picking other mirrors.
#
# If the mirrorlist does not work for you, as a fall back you can try the
# remarked out baseurl line instead.
#
#
[base]
name=CentOS-$releasever - Base - mirrors.aliyun.com
failovermethod=priority
baseurl=http://mirrors.aliyun.com/centos/$releasever/os/$basearch/
#http://mirrors.aliyun.com/centos/$releasever/os/$basearch/
#http://mirrors.cloud.aliyuncs.com/centos/$releasever/os/$basearch/
#gpgkey=http://mirrors.aliyun.com/centos/RPM-GPG-KEY-CentOS-7
#released updates
updates=
    
```

3. 界面选择后台手动配置选项



5.2. 进入Manager 8179安装页面显示需升级或安装失败后安装页面变为升级页面

【解决方法】

在服务端执行 yum remove transwarp-manager-common -y 后，重新进行Manager的安装。

5.3. 安装向导检查环境页面报错显示“无法解析hostname”

【问题描述】



【核心原因】

获取hostname 失败，需要开启机器的ping功能

【解决方法】

```
echo 0 >/proc/sys/net/ipv4/icmp_echo_ignore_all 开启
```

5.4. 安装向导显示提取文件失败，报错 “The package container-selinux conflicts with manager…”

【问题描述】

安装向导显示提取文件失败，报错 “The package container-selinux conflicts with manager…”

提取文件

提取文件失败

重试

```
The package container-selinux conflicts with Manager, please uninstall it and retry.
```

【核心原因】

container-selinux 冲突了

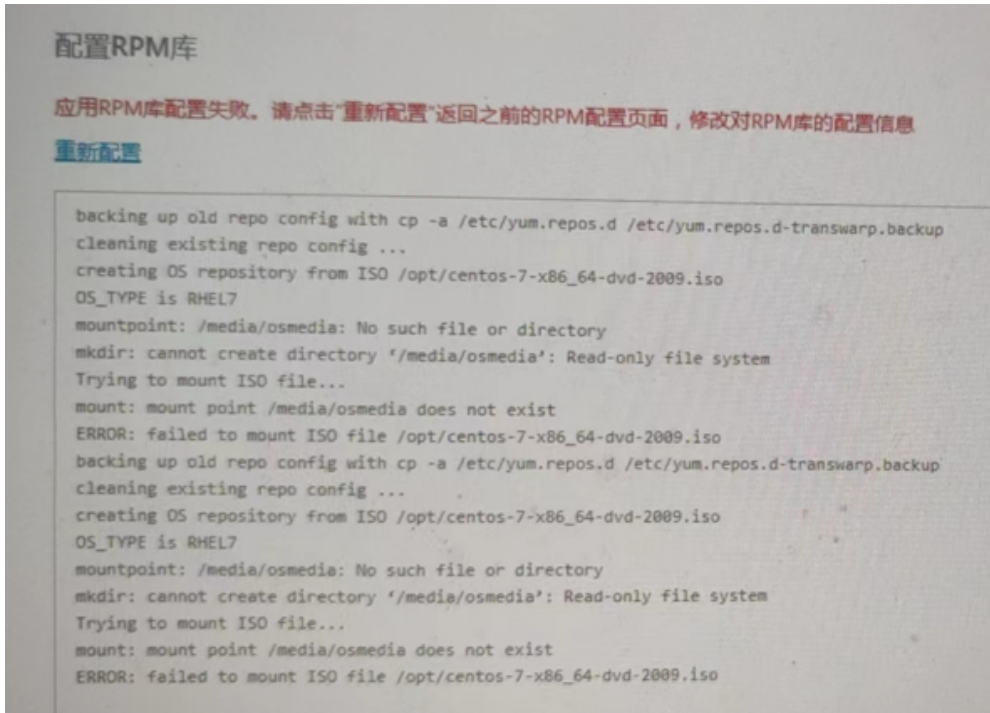
【解决方法】

执行下述命令：

```
rpm -qa|grep selinux
rpm -e container-selinux
```

注：可能会提示有依赖，需要逐步把依赖的包都卸载掉就可以了

5.5. 配置RPM库显示失败，报错显示“failed to mount ISO file /opt/xxx.iso”



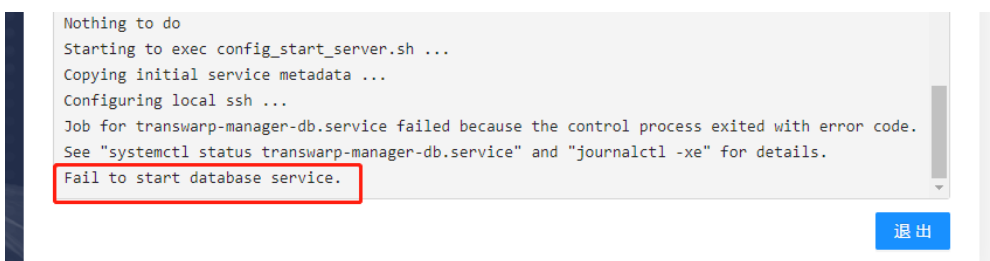
【核心原因】

/media/osmedia目录已经被挂载

【解决方法】

umount掉这个挂载再重试

5.6. 安装过程中报错显示“Fail to start database service”




```

[root@ XX ~]# systemctl status transwarp-manager-db.service
● transwarp-manager-db.service - Transwarp Manager database
   Loaded: loaded (/usr/lib/systemd/system/transwarp-manager-db.service; disabled; vendor preset: disabled)
   Active: failed (Result: exit-code) since Wed 2022-10-12 16:41:59 CST; 5min ago
   Process: 14089 ExecStartPre=/usr/lib/transwarp-manager/master/scripts/db_exec_start_pre.sh (code=exited, status=1/FAILURE)

Oct 12 16:41:59 [14089]: option to only install the mysql database and privilege tables.
Oct 12 16:41:59 [14089]: If you are using a binary release, you must either be at the top
Oct 12 16:41:59 XX [14089]: level of the extracted archive, or pass the --basedir option
Oct 12 16:41:59 [14089]: pointing to that location.
Oct 12 16:41:59 [14089]: The latest information about mysql_install_db is available at
Oct 12 16:41:59 [14089]: https://mariadb.com/kb/en/installing-system-tables-mysql_install_db
Oct 12 16:41:59 [1]: transwarp-manager-db.service: control process exited, code=exited status=1
Oct 12 16:41:59 XX [1]: Failed to start Transwarp Manager database.
Oct 12 16:41:59 [1]: Unit transwarp-manager-db.service entered failed state.
Oct 12 16:41:59 [1]: transwarp-manager-db.service failed.
[root@ XX ~]# journalctl -xe
Oct 12 16:48:04 sshd[23004]: Disconnected from 10.254.21.6 port 22986
Oct 12 16:48:04 XX sshd[23004]: pam_unix(sshd:session): session closed for user root
Oct 12 16:48:04 systemd-logind[948]: Removed session 132344.
-- Subject: Session 132344 has been terminated
-- Defined-By: systemd
-- Support: http://lists.freedesktop.org/mailman/listinfo/systemd-devel
-- Documentation: http://www.freedesktop.org/wiki/Software/systemd/multiseat
--
-- A session with the ID 132344 has been terminated.
Oct 12 16:48:07 pnode3 sshd[23068]: Accepted publickey for root from 10.254.21.6 port 23144 ssh2: RSA SHA256:1hVZRpFMdPvRnZD22pjQ
Oct 12 16:48:07 pnode3 systemd-logind[948]: New session 132345 of user root.
-- Subject: A new session 132345 has been created for user root
-- Defined-By: systemd
-- Support: http://lists.freedesktop.org/mailman/listinfo/systemd-devel
-- Documentation: http://www.freedesktop.org/wiki/Software/systemd/multiseat
--
-- A new session with the ID 132345 has been created for the user root.
--
-- The leading process of the session is 23068.

```

【解决方法】

用户需要修改创建数据库的参数:

/usr/lib/transwarp-manager/master/scripts/db_exec_start_pre.sh里的/usr/bin/mysql_install_db加一force参数。

5.7. 安装失败，报错显示 “The directory /var/lib/docker is not on a separate partition”



欢迎来到Manager安装向导

本向导将帮助您安装Manager。您可以在Manager上对所有集群和服务进行实时监控和管理。

- ✓ 阅读Java许可协议
- ✓ 检查环境
- ✓ 配置Manager
- ✓ 提取文件
- ✓ 配置RPM库
- 6 安装**

中文

安装

Manager安装失败

重试

```

Package psmisc-22.20-17.el7.x86_64 already installed and latest version
Nothing to do
start to install git-core
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
Package git-1.8.3.1-23.el7_8.x86_64 already installed and latest version
Nothing to do
start to install mariadb-server
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
Package 1:mariadb-server-5.5.68-1.el7.x86_64 already installed and latest version
Nothing to do
start to install openssh-server openssh-clients
Loaded plugins: fastestmirror
Loading mirror speeds from cached hostfile
Package openssh-server-7.4p1-21.el7.x86_64 already installed and latest version
Package openssh-clients-7.4p1-21.el7.x86_64 already installed and latest version
Nothing to do
Starting to exec config_start_server.sh ...
device was mounted to /var/lib/docker
The directory /var/lib/docker is not on a separate partition.

```

The directory /var/lib/docker is not on a separate partition.

退出

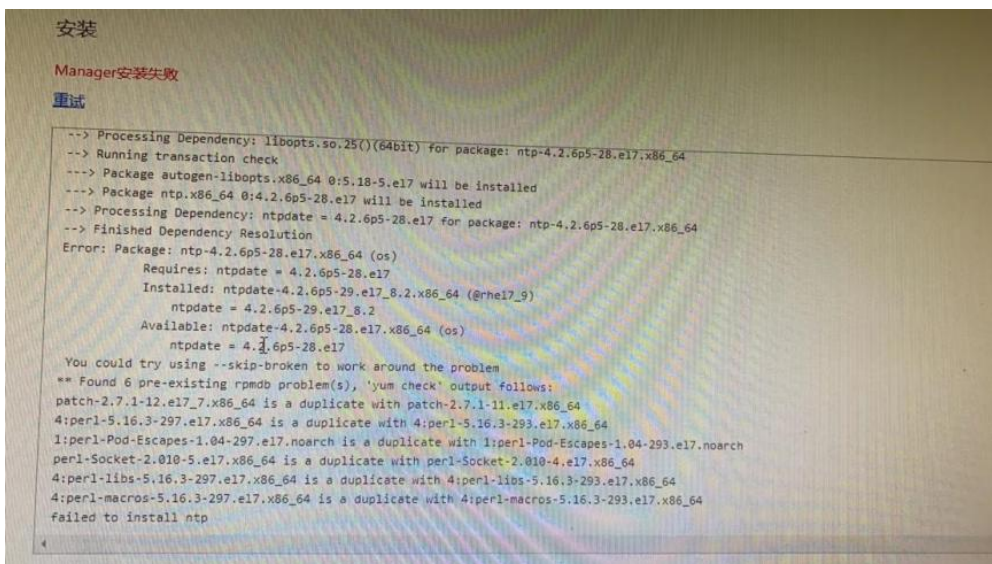
【报错原因】

显示不是一个独立的分区，没有挂载docker目录

【解决方法】

可以执行 `mount | grep "/var/lib/docker"` 检查，如果未挂载，需要把docker目录重新挂载，挂载后重试即可。如果重启后没有自动挂载上，依然失败，修改/etc/fstab后重启，再次重新安装。

5.8. 安装失败，报错显示“failed to install ntp”



```

安装
Manager安装失败
重试
--> Processing Dependency: libopts.so.25()(64bit) for package: ntp-4.2.6p5-28.el7.x86_64
--> Running transaction check
--> Package autogen-libopts.x86_64 0:5.18-5.el7 will be installed
--> Package ntp.x86_64 0:4.2.6p5-28.el7 will be installed
--> Processing Dependency: ntpdate = 4.2.6p5-28.el7 for package: ntp-4.2.6p5-28.el7.x86_64
--> Finished Dependency Resolution
Error: Package: ntp-4.2.6p5-28.el7.x86_64 (os)
Requires: ntpdate = 4.2.6p5-28.el7
Installed: ntpdate-4.2.6p5-29.el7_8.2.x86_64 (@rhel7_9)
           ntpdate = 4.2.6p5-29.el7_8.2
Available: ntpdate-4.2.6p5-28.el7.x86_64 (os)
           ntpdate = 4.2.6p5-28.el7
You could try using --skip-broken to work around the problem
** Found 6 pre-existing rpmdb problem(s), 'yum check' output follows:
patch-2.7.1-12.el7_7.x86_64 is a duplicate with patch-2.7.1-11.el7.x86_64
4:perl-5.16.3-297.el7.x86_64 is a duplicate with 4:perl-5.16.3-293.el7.x86_64
1:perl-Pod-Escapes-1.04-297.el7.noarch is a duplicate with 1:perl-Pod-Escapes-1.04-293.el7.noarch
perl-Socket-2.010-5.el7.x86_64 is a duplicate with perl-Socket-2.010-4.el7.x86_64
4:perl-libs-5.16.3-297.el7.x86_64 is a duplicate with 4:perl-libs-5.16.3-293.el7.x86_64
4:perl-macros-5.16.3-297.el7.x86_64 is a duplicate with 4:perl-macros-5.16.3-293.el7.x86_64
failed to install ntp

```

【核心原因】

设置的操作系统软件源有冲突，如果用户自己配了软件源，可以在安装时直接选择后台手工配置，就不用manager配了。

【解决方法】

把已有的ntp卸载后重试即可。

5.9. 安装失败，报错显示“系统已经安装过Manager-8.1.3-1.noarch”，但服务连不上

【问题描述】

安装社区版的时候，没有创建 /var/lib/docker 目录，导致最后一步安装失败，安装向导显示“系统已经安装过Manager-8.1.3-1.noarch”，连不上服务



```

2022-10-12 16:30:28.033 INFO 20120 --- [pinned-executor-0] i.t.m.installer.core.PrepareStepTask : prepare result of manager-check was N
2022-10-12 16:31:13.121 INFO 20120 --- [pinned-executor-0] i.t.m.installer.core.ProcessSettingTask : processing manager-check
2022-10-12 16:31:13.122 ERROR 20120 --- [pinned-executor-0] i.t.m.installer.core.ProcessSettingTask : error processing manager-check

java.lang.NullPointerException: null
    at io.transwarp.manager.installer.step.ManagerCheckStep.process(ManagerCheckStep.java:64)
    at io.transwarp.manager.installer.core.ProcessSettingTask.run(ProcessSettingTask.java:29)
    at java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:511)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:750)

2022-10-12 16:31:14.536 INFO 20120 --- [pinned-executor-0] i.t.m.installer.core.PrepareStepTask : prepare result of manager-check was N

```

【解决方法】

在服务端执行

```
rpm -e transwarp-manager-common
rpm -e transwarp-manager
```

或yum remove transwarp-manager-common -y重新安装即可

5.10. 安装失败，报错显示“failed to install MariaDB server”

【报错信息】

导

manager

安装

Manager安装失败

[重试](#)

```

start to install openssh-server openssh-clients
已加载插件: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
软件包 openssh-server-7.4p1-21.el7.x86_64 已安装并且是最新版本
软件包 openssh-clients-7.4p1-21.el7.x86_64 已安装并且是最新版本
无须任何处理
start to install java-1.8.0-openjdk-devel
已加载插件: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
软件包 1:java-1.8.0-openjdk-devel-1.8.0.262.b10-1.el7.x86_64 已安装并且是最新版本
无须任何处理
start to install fontconfig
已加载插件: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
软件包 fontconfig-2.13.0-4.3.el7.x86_64 已安装并且是最新版本
无须任何处理
start to install transwarp-manager
已加载插件: fastestmirror, langpacks
Loading mirror speeds from cached hostfile
软件包 transwarp-manager-9.3.1-1.noarch 已安装并且是最新版本
无须任何处理
Starting to exec config_start_server.sh ...
failed to install MariaDB Server

```

【核心原因】

显示与机器上已有的mysql冲突

```

---> 软件包 mariadb-libs.x86_64.1.5.5.68-1.el7 将被 安装
Removing mariadb.x86_64 1:5.5.68-1.el7 - u due to obsoletes from
installed mysql-community-client-5.7.13-1.el7.x86_64
Removing mariadb-libs.x86_64 1:5.5.68-1.el7 - u due to obsoletes
from installed mysql-community-libs-5.7.13-1.el7.x86_64
Removing mariadb-server.x86_64 1:5.5.68-1.el7 - u due to obsoletes
from installed mysql-community-server-5.7.13-1.el7.x86_64
--> 正在使用新的信息重新解决依赖关系

```

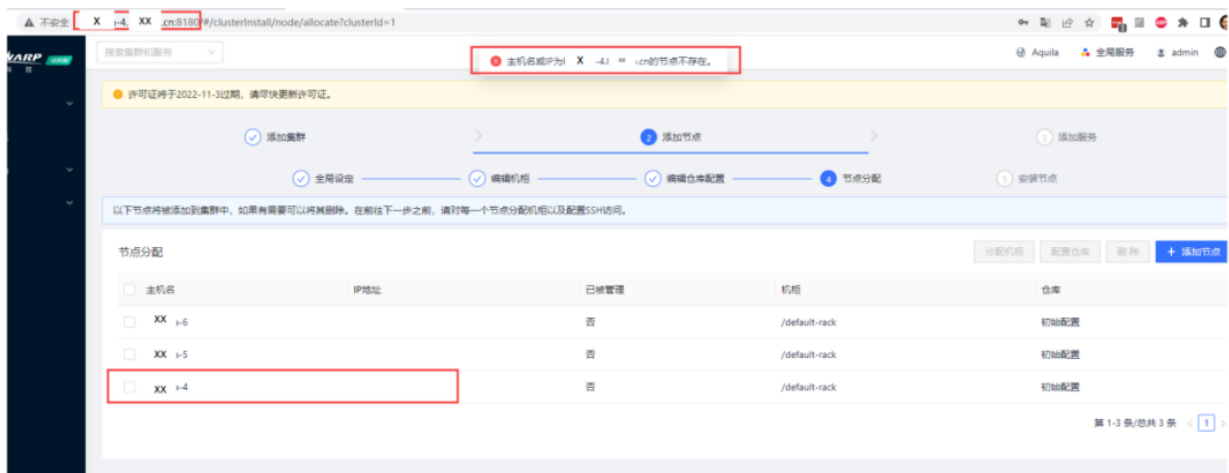
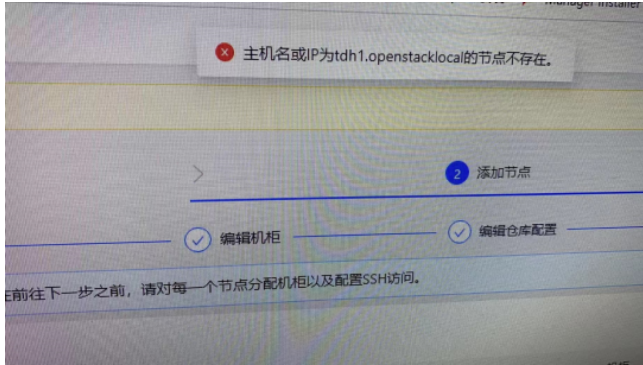
【解决方法】

建议删除机器上的mysql，如果用户不想删除，可以考虑换台机器装Manager。

6. 配置集群过程中可能遇到的问题

6.1. 添加节点过程中报错显示“主机名或IP为xxx的节点不存在”

【报错信息】



【核心原因】

出现此类报错，请优先检查hostname以及/etc/hosts文件中的主机名是否一致。

下面将介绍两个因为同一个原因导致报错的用户示例，出现报错的核心原因是搞混了 hostname 和带domain的hostname。

注意，hostname和全域名（FQDN）是不一样的，FQDN包含两个部分，hostname和域名，hostname就是FQDN的第一个点前面的部分，所以hostname是不包含点的。

如果您也遇到了类似的问题，可参考下述方法进行解决。

6.1.1. 用户现象1

用户设置的hostname如下图所示，此时ping进行解析的时候实际上解析的是前半部分，比如tdh2而不是tdh2.openstacklocal

```

-rw-r--r-- 1 root root 1845710488 7月 27 10:49 TDH-Platfo
drwxr-xr-x 3 root root 77 7月 27 10:45 TDH-Platfor
61 5月 4 10:44 transwarp-8
[root@tdh1 home]# cat /ect/hosts
cat: /ect/hosts: 没有那个文件或目录
[root@tdh1 home]# cd /
[root@tdh1 /]# cat /etc/hosts
127.0.0.1 localhost localhost.localdomain localhost4 localhost4
::1 localhost.localdomain localhost6 localhost6.localdomain6
10.192.192.192 tdh1 tdh1.openstacklocal
10.193.193.193 tdh2 tdh2.openstacklocal
10.194.194.194 tdh3 tdh2.openstacklocal
[root@tdh1 /]#

```

【解决方法】

修改hostname后修改hosts文件即可解决（修改为不带后缀“.”），需要注意，修改完之后需要用systemctl restart transwarp-manager-agent重启agent。

6.1.2. 用户现象2

用户的hostname设置为了xx-4.bigdata.cn，但是host文件设置的内容是：

```

10.xxxx.2 bigdata-6
10.xxxx.3 bigdata-4
10.xxxx.4 bigdata-5

```

不一致导致报错

【解决方法】

使用下述指令修改hostname：

```
hostnamectl set-hostname bigdata-4
```

其他节点同理，请务必确保hostname以及/etc/hosts文件中的主机名一致

6.2. 安装集群xx中的3个节点失败，日志报错显示：**error.nodeHostNotFound**

【报错信息】

```

2023-05-08T15:28:40.366 [SSH] execute command: test -e /etc/yum.repos.d-transwarp.backup || cp -a /etc/yum.repos.d /etc/yum.repos.d-transwarp.backup
2023-05-08T15:28:40.156 [SSH] exit status: 0
2023-05-08T15:28:40.158 [SSH] execute command: rm -rf /etc/yum.repos.d/*.repo
2023-05-08T15:28:40.366 [SSH] exit status: 0
2023-05-08T15:28:40.370 [Master] The Task 13 run failed: java.lang.RuntimeException: io.transwarp.manager.master.common.AppException: error.nodeHostNotFound
at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask(TaskDriver.java:226)
at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$$FastClassBySpringCGLIB$$e353057f.invoke(<generated>)
at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:783)
at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:753)
at org.springframework.aop.interceptor.AsyncExecutionInterceptor.Lambda$invoke$0(AsyncExecutionInterceptor.java:115)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: io.transwarp.manager.master.common.AppException: error.nodeHostNotFound
at io.transwarp.manager.master.service.NodeService.getNodeByIp(NodeService.java:284)
at io.transwarp.manager.master.service.NodeService$$FastClassBySpringCGLIB$$8d6ce278c.invoke(<generated>)
at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
at org.springframework.aop.framework.CglibAopProxy$DynamicAdvisedInterceptor.intercept(CglibAopProxy.java:688)
at io.transwarp.manager.master.service.NodeService$$EnhancerBySpringCGLIB$$833cf59f.getNodeByIp(<generated>)
at java.util.stream.ReferencePipeline$3$1.accept(ReferencePipeline.java:193)
at java.util.ArrayList$Itr.forEachRemaining(ArrayList.java:899)
at java.util.Spliterators$IteratorSpliterator.forEachRemaining(Spliterators.java:1801)
at java.util.stream.AbstractPipeline.copyInto(AbstractPipeline.java:481)
at java.util.stream.AbstractPipeline.wrapAndCopyInto(AbstractPipeline.java:471)
at java.util.stream.ReduceOps$ReduceOp.evaluateSequential(ReduceOps.java:708)
at java.util.stream.AbstractPipeline.evaluate(AbstractPipeline.java:234)
at java.util.stream.ReferencePipeline.collect(ReferencePipeline.java:499)
at org.jooq.Lambda.SeqImpl.collect(SeqImpl.java:190)
at org.jooq.Lambda.Seq.toList(Seq.java:9194)
at org.jooq.Lambda.SeqImpl.toList(SeqImpl.java:558)
at io.transwarp.manager.master.service.wrapper.ManagerWrapper.getManagers(ManagerWrapper.java:70)
at io.transwarp.manager.master.operation.execution.Localrunner.NodeInstallTaskLocalRunner.appointYumRepos(NodeInstallTaskLocalRunner.java:323)
at io.transwarp.manager.master.operation.execution.Localrunner.NodeInstallTaskLocalRunner.execute(NodeInstallTaskLocalRunner.java:93)
at io.transwarp.manager.master.operation.execution.Localrunner.NodeInstallTaskLocalRunner.execute(NodeInstallTaskLocalRunner.java:43)
at io.transwarp.manager.master.operation.execution.Localrunner.AbstractTaskLocalRunner.run(AbstractTaskLocalRunner.java:98)

```

【核心原因】

- Host文件配置有问题

用户的3个节点名称为别为tdh1, tdh2, tdh3, 但是/etc/hosts 文件的第一行必须为127.0.0.1的记录, 不能将此行注释掉。

```

[root@tdh1 ~]# ping tdh2 -c 4
PING tdh2 (10.0.12.12) 56(84) bytes of data:
64 bytes from tdh2 (10.0.12.12): icmp_seq=1 ttl=64 time=0.156 ms
64 bytes from tdh2 (10.0.12.12): icmp_seq=2 ttl=64 time=0.167 ms
64 bytes from tdh2 (10.0.12.12): icmp_seq=3 ttl=64 time=0.161 ms
64 bytes from tdh2 (10.0.12.12): icmp_seq=4 ttl=64 time=0.191 ms

--- tdh2 ping statistics ---
4 packets transmitted, 4 received, 0% packet loss, time 2999ms
rtt min/avg/max/mdev = 0.156/0.168/0.191/0.020 ms
[root@tdh1 ~]# cat /etc/hosts
127.0.0.1 tdh1

::1 tdh1

10.0.12.4 tdh1
10.0.12.12 tdh2
10.0.12.10 tdh3
[root@tdh1 ~]#

```

【解决方法】

需要参考下方黄色框里的内容, 与其基本一致, 部署时不要修改这部分的内容, 不能把当前主机名写在该行中。

```

nameserver 114.114.114.114
[root@idc8 inceptor1]# hostname
idc8
[root@idc8 inceptor1]# cat /etc/hosts
127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6
192.168.111.9 idc9
192.168.111.8 idc8
192.168.111.12 idc12
192.168.111.11 idc11
192.168.111.114 idcv14
172.18.124.32 idc20

```

6.3. 添加节点步骤报错显示“节点无法加入，原因：error.nodeGetHostnameFail”

【报错信息】



【核心原因】

非root用户，必须使用无密码的sudo用户

```

Tasks = 2, completed tasks = 1315
2023-05-24 10:40:55.930 INFO [http-nio-8180-exec-6] n.s.sshj.transport.random.JCERandom: Creating new SecureRandom.
2023-05-24 10:40:55.935 INFO [http-nio-8180-exec-6] net.schmizz.sshj.transport.TransportImpl: Client identity string: SSH-2.0-SSHJ_0.32.0
2023-05-24 10:40:55.949 INFO [http-nio-8180-exec-6] net.schmizz.sshj.transport.TransportImpl: Server identity string: SSH-2.0-OpenSSH_7.4
2023-05-24 10:40:56.197 INFO [http-nio-8180-exec-6] net.schmizz.sshj.transport.TransportImpl: Disconnected - BY_APPLICATION
2023-05-24 10:40:56.199 ERROR [http-nio-8180-exec-6] i.t.manager.master.node.NodeService: Can't get hostname on 192.168.111.11:31; sudo: 需要密码
2023-05-24 10:40:56.213 INFO [http-nio-8180-exec-6] i.t.manager.common.util.HostsFileUtils: /etc/hosts on manager node updated!
2023-05-24 10:40:56.214 INFO [http-nio-8180-exec-6] i.t.manager.common.util.HostsFileUtils: Manager addressCache updated!
2023-05-24 10:44:15.674 INFO [http-nio-8180-exec-7] n.s.sshj.transport.random.JCERandom: Creating new SecureRandom.
2023-05-24 10:44:15.678 INFO [http-nio-8180-exec-7] net.schmizz.sshj.transport.TransportImpl: Client identity string: SSH-2.0-SSHJ_0.32.0
2023-05-24 10:44:15.688 INFO [http-nio-8180-exec-7] net.schmizz.sshj.transport.TransportImpl: Server identity string: SSH-2.0-OpenSSH_7.4
2023-05-24 10:44:15.958 INFO [http-nio-8180-exec-7] net.schmizz.sshj.transport.TransportImpl: Disconnected - BY_APPLICATION
2023-05-24 10:44:15.960 ERROR [http-nio-8180-exec-7] i.t.manager.master.node.NodeService: Can't get hostname on 192.168.111.11:31; sudo: 需要密码
2023-05-24 10:44:15.989 INFO [http-nio-8180-exec-7] i.t.manager.common.util.HostsFileUtils: /etc/hosts on manager node updated!

```

6.4. 添加节点时，访问节点身份使用sudo用户，报错显示“节点无法加入，sudo:需要密码”



【用户描述】

- 在安装社区版的时候，用的是非root用户，但是有sudo权限，这里提示需要sudo密码，这个应该在哪儿设置？

【解决方法】

- 需要配置下，使sudo用户无需输入密码。具体执行命令如下：

1. root用户登录，执行 `vim /etc/sudoers`
2. 修改 `/etc/sudoers`文件

```
%admin ALL= (ALL) 改为  %admin ALL= (ALL) NOPASSWD: ALL
```

3. 保存并退出

重新进行添加节点操作即可。

7. 安装产品服务过程中可能遇到的问题

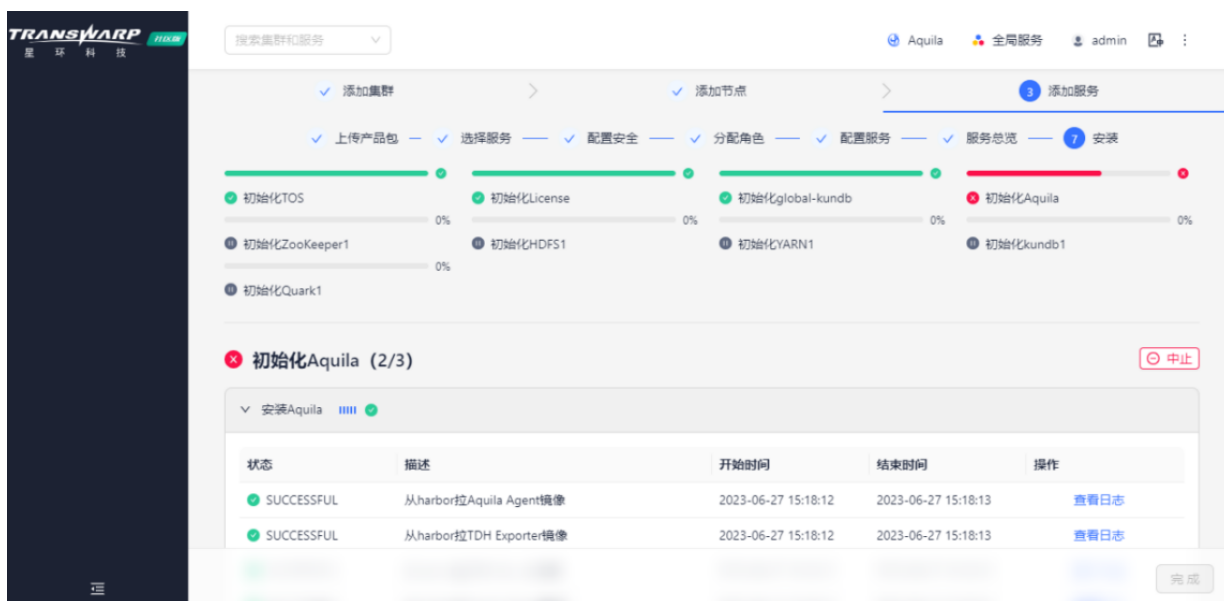
7.1. 通用的排错方法

7.1.1. 添加服务阶段初始化服务失败如何定位问题

7.1.1.1. 示例一：初始化Aquila失败，报错步骤为“通过TOS启动service失败”

【问题描述】

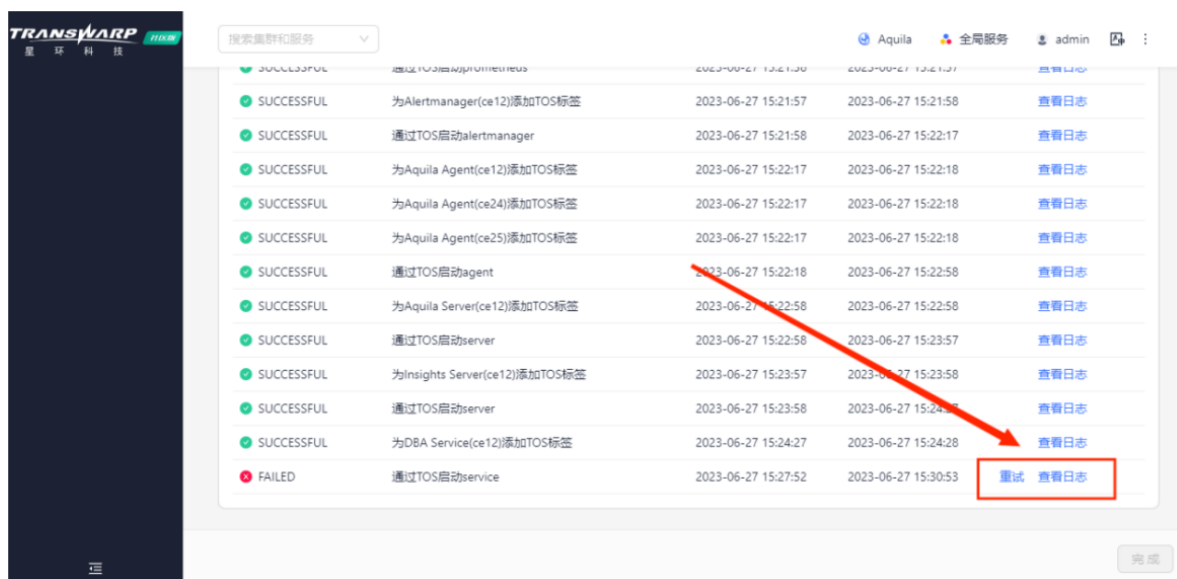
社区版安装时，在添加服务步骤初始化Aquila失败，报错内容为“通过TOS启动service失败”，最终导致添加服务失败。显示内容如下图所示：



【排查过程】

1. 查看报错日志

通常在安装过程中出现报错，可以优先查看该步骤右侧的日志排查报错原因。




```

2023-06-27T15:27:52.608017 [Master] ===== Task 157 start to run. =====
2023-06-27T15:27:52.610558 [Master] Starting task local part ...
2023-06-27T15:27:52.613746 [Master] Start handle kubectl task
2023-06-27T15:27:52.617968 [Master] Target yaml path on manager is /var/lib/transwarp-manager/master/content/
resources/services/aquila/dba-service.yaml
2023-06-27T15:27:52.618006 [Master] Start to generate dba-service.yaml on manager...
2023-06-27T15:27:52.707673 [Master] generated dba-service.yaml on [Manager]
2023-06-27T15:27:52.960353 [Master] Start [Create Roles] ...
2023-06-27T15:27:52.960405 [Master] Start executing [kubectl --server=https://127.0.0.1:6443 --certificate-
authority=/srv/kubernetes/ca.pem --client-certificate=/srv/kubernetes/admin.pem --client-key=/srv/kubernetes/
admin-key.pem scale --replicas=1 -f /var/lib/transwarp-manager/master/content/resources/services/aquila/dba-
service.yaml]
2023-06-27T15:27:53.181953 [Master] [Create Roles] success.
2023-06-27T15:27:53.182014 [Master] Task local part ended.
2023-06-27T15:27:53.182036 [Master] Starting task remote part ...
2023-06-27T15:27:53.186043 [Master] Waiting 1 DBA_SERVICE role(s) in Aquila to become Healthy within 600 s
2023-06-27T15:30:53.272493 [Master] Latest health check result of roles:
DAEMON_CHECK DOWN at 2023-06-27T15:30:52.632
DBA_SERVICE has no Pod on ce12VITAL_SIGN_CHECK DOWN at 2023-06-27T15:30:47.836
DBA_SERVICE has no ready pod on ce12
2023-06-27T15:30:53.272005 [Master] fail to run task remote part:
io.transwarp.manager.master.operation.execution.entity.TaskDownException: java.lang.IllegalStateException: 1
DBA_SERVICE role(s) in Aquila didn't become healthy within 180 s
at
io.transwarp.manager.master.operation.execution.localrunner.KubectlTaskLocalRunner.postRemote(KubectlTaskLocalRunn
er.java:898)
at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask(TaskDriver.java:287)
at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$$FastClassBySpringCGLIB$
$e353057f.invoke(<generated>)
at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
at

```

由此可知，安装任务失败是由于Aquila的角色之一DBA Service的Pod出现问题导致，所以需要进一步查看Pod的状态以及日志信息，来定位根因。

2. 查询DBA Service Pod状态

确定失败服务后，使用 `kubectl get pod | grep <组件名称>` 命令查询组件相关Pod的运行状态。由下图可以看出，DBA Service Pod处于pending状态。因此，接下来需要展开该Pod的详细信息，来进一步定位问题。

Tips: 如果您在查看Pod状态时，显示为Failed或其他状态比如冲突，您可以通过 `kubectl logs <pod id>`，来查看Pod的相关日志。

```

Last login: Tue Jun 27 15:52:08 2023 from 172.16.221.61
[root@ce12 ~]# kubectl get pod | grep aquila
alertmanager-aquila-5c745b4cfb-5929t          1/1      Running    0          47m
aquila-agent-aquila-57d6cbbc96-9b12f        1/1      Running    0          47m
aquila-agent-aquila-57d6cbbc96-bn74q        1/1      Running    0          47m
aquila-agent-aquila-57d6cbbc96-r85cn        1/1      Running    0          47m
aquila-server-aquila-84dd5565fd-rlh9m       1/1      Running    0          46m
dba-service-aquila-5f44b6846f-tg797        0/1      Pending   0          45m
insights-server-aquila-6bdc69cff7-vfrz8     1/1      Running    0          45m
kube-state-metrics-aquila-f876fffb5-9cj55   1/1      Running    0          48m
prometheus-aquila-777799976-2jbwq          1/1      Running    0          47m
tdh-exporter-aquila-5b85fdf9f5-7rsgz       1/1      Running    0          48m
tdh-exporter-aquila-5b85fdf9f5-gtc8m       1/1      Running    0          48m
tdh-exporter-aquila-5b85fdf9f5-hg976      1/1      Running    0          48m
[root@ce12 ~]# kubectl logs dba-service-aquila-5f44b6846f-tg797
[root@ce12 ~]# █

```

3. 查看Pod信息

使用命令 `kubectl describe pod <pod id>` 查看信息。

```
[root@ce12 ~]# kubectl describe pod dba-service-aquila-5f44b6846f-tg797
Name: dba-service-aquila-5f44b6846f-tg797
Namespace: default
Priority: 0
Priority Class Name: low-priority
Node: <none>
Labels: name=dba-service-aquila
pod-template-hash=5f44b6846f
podConflictName=dba-service-aquila
Annotations: serviceInstanceName: Aquila
Status: Pending
IP:
IPs: <none>
Controlled By: ReplicaSet/dba-service-aquila-5f44b6846f
Containers:
  dba-service-aquila:
    Image: transwarp/aquila:tdh_aquila-9.3-ce-final
    Port: <none>
    Host Port: <none>
    Args:
      boot.sh
      DBA_SERVICE
    Limits:
      cpu: 1
      memory: 1Gi
    Requests:
      cpu: 1
      memory: 1Gi
    Readiness: http-get https://:4040/ delay=10s timeout=1s period=10s #success=1 #failure=3
    Environment:
      WATCHMAN_CONF: /etc/aquila/conf/dbaservice/application.yml
      SPRING_DATASOURCE_PASSWORD: password
      SERVER_SSL_KEY-STORE-PASSWORD: changeit
      MANAGER_DBA_LOGIN_PASSWORD:
      LVT_MSG_PORT: 60606
      MANAGER_AQUILA_LOGIN_PASSWORD: W0gIFQoqJfF89
      RUNNING_MODE: dbaservice
      TRANSWARP_NETWORK_MODE: HOST
      TRANSWARP_NETWORK_TYPE: HOST
    Mounts:
      /etc/aquila/conf from conf (rw)
      /etc/global-kundb/conf from global-kundb (rw)
      /etc/localtime from timezone (rw)
      /etc/tos/conf from tos (rw)
      /etc/transwarp/conf from transwarphosts (rw)
      /opt/kubernetes from kubelet-exec (rw)
      /srv/kubernetes from kubeconfig (rw)
      /usr/lib/transwarp/plugins from plugin (rw)
      /var/log/aquila/ from log (rw)
      /var/run/secrets/kubernetes.io/serviceaccount from default-token-dltfg (ro)
      /vdir from mountbind (rw)
Conditions:
```

```
tos:
  Type: HostPath (bare host directory volume)
  Path: /etc/tos/conf
  HostPathType:
global-kundb:
  Type: HostPath (bare host directory volume)
  Path: /etc/global-kundb/conf
  HostPathType:
default-token-dltfg:
  Type: Secret (a volume populated by a Secret)
  SecretName: default-token-dltfg
  Optional: false
QoS Class: Guaranteed
Node-Selectors: dba-service-aquila=true
Tolerations:
  node.kubernetes.io/not-ready:NoExecute for 1800s
  node.kubernetes.io/unreachable:NoExecute for 1800s
Events:
  Type Reason Age From Message
  ----
Warning FailedScheduling <unknown> default-scheduler 0/3 nodes are available: 1 Insufficient cpu, 2 node(s) didn't match node selector.
Warning FailedScheduling <unknown> default-scheduler 0/3 nodes are available: 1 Insufficient cpu, 2 node(s) didn't match node selector.
[root@ce12 ~]#
```

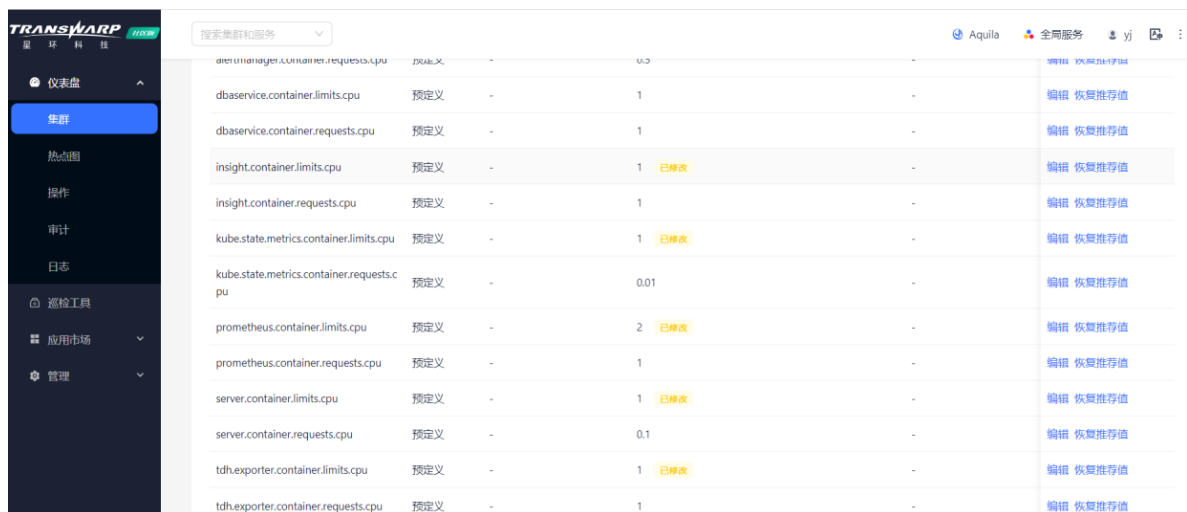
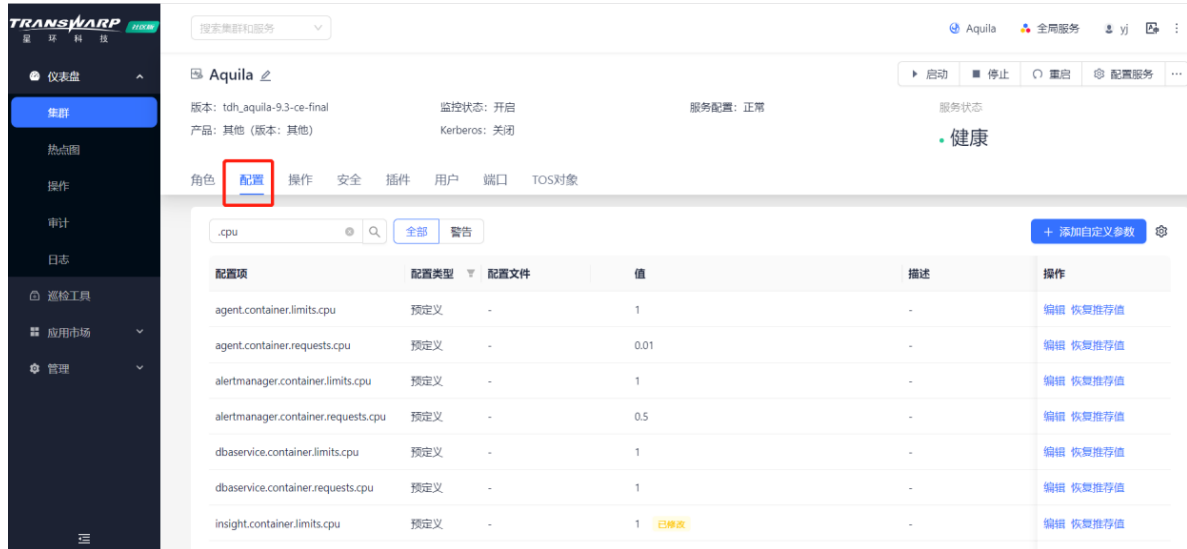
由上图可以看出，这个问题的根因为服务器分配给Aquila的资源不足，CPU资源限制导致该Pod无法启动。

至此，该问题报错的根因已找到，接下来就是做出问题修复，在这提供解决思路和预防措施。

【解决方法】

1. 解决思路：修改Aquila服务的资源分配

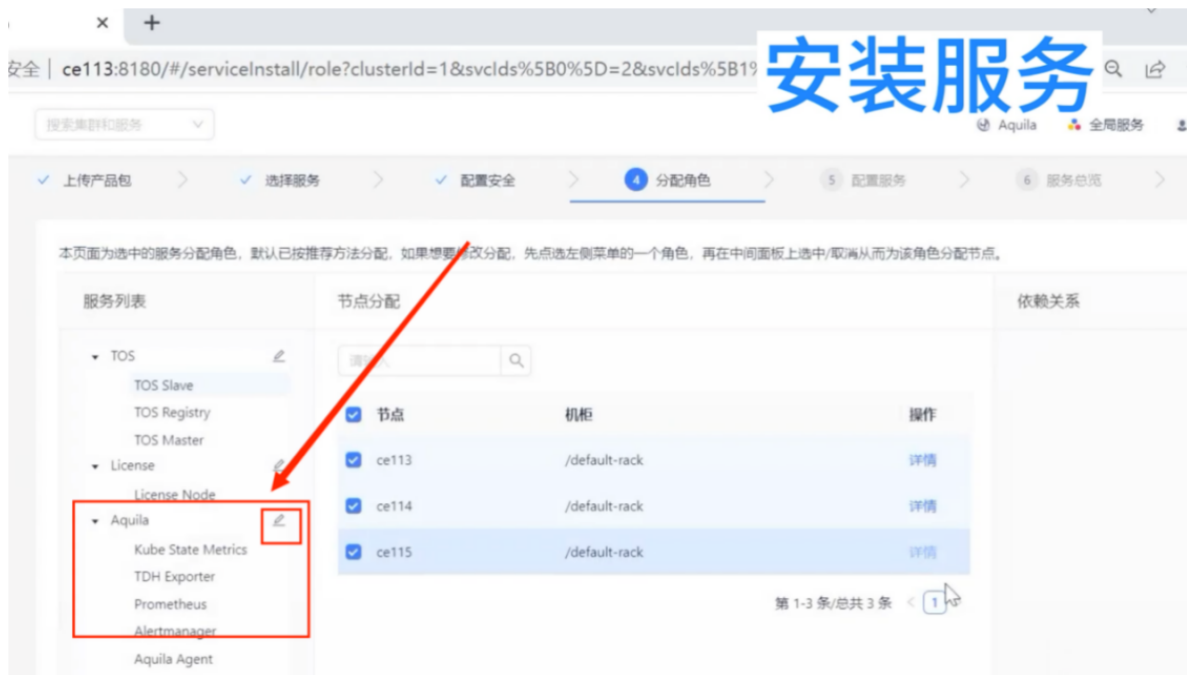
TDH基于TCOS云原生统一资源调度系统可以实现对各服务所占用的硬件、软件资源做出统一调配。因此在Manager界面可以配置分配给服务的CPU资源，以实现服务的正常运行。修改完成后需要点击“配置服务”，然后重启服务。参考示例如下：



2. 预防措施：均匀分配Aquila角色安装节点

在TDH社区版中，会为大家提供默认的服务配置，大部分服务角色会默认安装在主节点中。若您的服务器资源较为紧张，您可以通过在安装阶段将Aquila个别服务角色均匀分配在不同节点上，缓解单一节点的负载压力，来预防该问题的发生。

服务配置所处的阶段为：集群搭建—安装服务—分配角色（参考下图）。点击角色可选择不同的安装节点，选择其他节点即可。



7.1.1.2. 示例二：初始化Quark失败，报错步骤为“无法启动Quark的角色”



【核心原因】

- 数据库中表缺失

因为某些操作，导致数据库中缺失“INCEPTOR_LIBRARY”表，通过执行SQL命令创建表实现修复。

【排查过程】

1. 查看Manager界面报错时右侧的查看日志，检查报错原因。

一个小tips：通常页面显示某一步骤失败，可以查看失败日志，日志里一般会提示是执行了什么命令，然后等待600s没变成健康状态，比如这个问题所对应的报错日志中有这么一个信息。

```

2023-06-26T03:45:46.646290 [Master] ===== Task 116 start to run. =====
2023-06-26T03:45:46.646882 [Master] Starting task local part ...
2023-06-26T03:45:46.647670 [Master] Start handle role task...
2023-06-26T03:45:46.709894 [Master] execute command: DirectiveDetail.RawShellO (script=systemctl start transwarp-quark-metastore@quark1)
2023-06-26T03:45:46.709950 [Master] Task local part ended.
2023-06-26T03:45:46.709975 [Master] Starting task remote part ...
2023-06-26T03:45:46.711789 [Agent] Execute command: systemctl start transwarp-quark-metastore@quark1
2023-06-26T03:45:47.339305 [Master] Waiting Quark Metastore (Quark1.tw-node2125) to become Healthy within 600 s
2023-06-26T03:48:47.386561 [Master] Latest health check result of roles:
DAEMON_CHECK DOWN at 2023-06-26T03:48:44.223

The INCEPTOR_METASTORE on node tw-node2125 was unhealthy.

2023-06-26T03:48:47.386631 [Master] Fail to run task remote part: java.lang.IllegalStateException: Quark Metastore (Quark1.tw-node2125) didn't become healthy within 180 s
    at io.transwarp.manager.master.operation.execution.localrunner.AbstractTaskLocalRunner.waitRolesHealthy(AbstractTaskLocalRunner.java:721)
    at io.transwarp.manager.master.operation.execution.localrunner.RoleTaskLocalRunner.postRemote(RoleTaskLocalRunner.java:255)
    at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask(TaskDriver.java:287)
    at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$$FastClassBySpringCGLIB$$e353057f.invoke(<generated>)
    at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
    at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:793)
    at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
    at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:763)
    at org.springframework.aop.interceptor.AsyncExecutionInterceptor.lambda$invoke$0(AsyncExecutionInterceptor.java:115)
    at java.base/java.util.concurrent.FutureTask.run(FutureTask.java:264)
    at java.base/java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1128)
    at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628)
    at java.base/java.lang.Thread.run(Thread.java:829)
Caused by: org.awaitility.core.ConditionTimeoutException: still DOWN within 180 s
    at io.transwarp.manager.master.operation.execution.localrunner.AbstractTaskLocalRunner.lambda$waitRolesHealthy$7(AbstractTaskLocalRunner.java:700)
    at org.awaitility.core.CallableCondition$ConditionEvaluationWrapper.eval(CallableCondition.java:99)
    at org.awaitility.core.ConditionAwaiter$ConditionPoller.call(ConditionAwaiter.java:248)
    at org.awaitility.core.ConditionAwaiter$ConditionPoller.call(ConditionAwaiter.java:235)
    ... 4 more

```

所以根据日志所提示的内容进一步排错，查看transwarp-quark-metastore@quark1该服务的状态是否有问题才导致start的时候失败。

2. 查看服务状态

执行docker ps -a命令查看容器id后，执行docker exec -it <containerid> bash的命令，进入容器。进入容器后执行systemctl status transwarp-quark-metastore@quark1 命令查看服务状态

```

[root@tw-node2125 ~/banber_deploy/docker-compose/banber_app]# docker exec -it 6d627423110d bash
[root@tw-node2125 /]# systemctl status transwarp-quark-metastore@quark1
transwarp-quark-metastore@quark1.service - Transwarp QUARK METASTORE
   Loaded: loaded (/usr/lib/systemd/system/transwarp-quark-metastore@.service, disabled)
   Active: failed (failed)

```

服务状态显示为失败，因此执行systemctl start transwarp-quark-metastore@quark1启动服务。

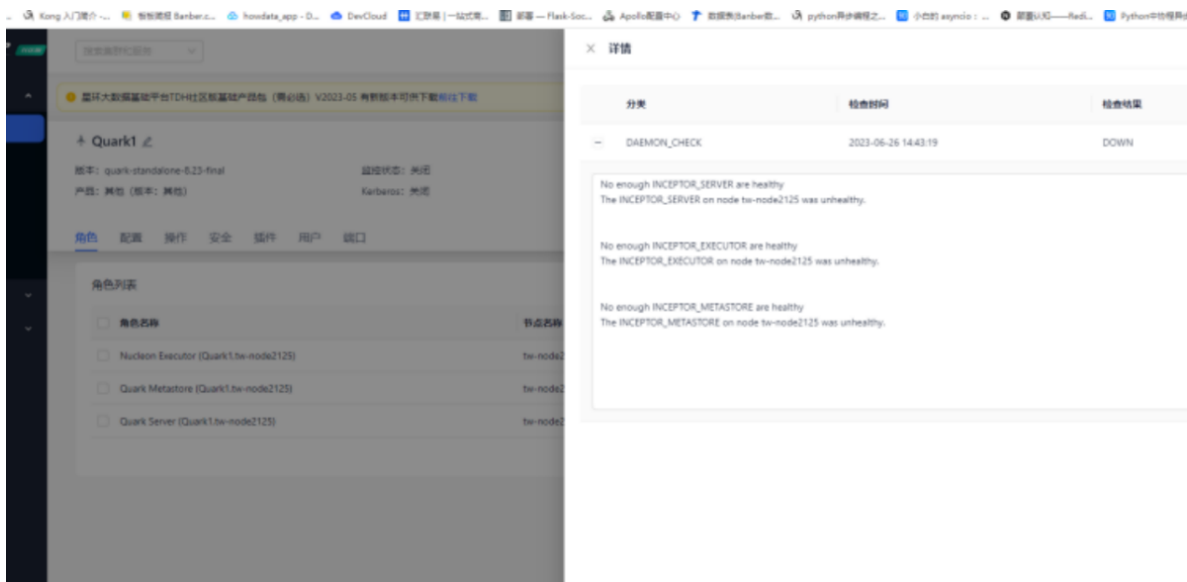
```

[root@tw-node2125 /]# systemctl status transwarp-quark-metastore@quark1
transwarp-quark-metastore@quark1.service - Transwarp QUARK METASTORE
   Loaded: loaded (/usr/lib/systemd/system/transwarp-quark-metastore@.service, disabled)
   Active: active (running)

```

3. 修复后重试

执行完成后，在Manger界面重试启动Quark，仍报错。



4. 继续排查错误

进入容器后用 `journalctl -u` 查看 `metastore` 角色的信息，执行 `journalctl -u transwarp-quark-metastore@quark1`

```

+ port=3308
+ mysql -h 127.0.0.1 -P 3308 -u inceptoruser --password=password metastore quark1 -e 'select * from INCEPTOR_LIBRARY'
ERROR 1146 (42S02) at line 1: Table 'metastore_quark1.INCEPTOR_LIBRARY' doesn't exist
+ return 1
+ echo 'Not ready, wait for 5 seconds ...'
Not ready, wait for 5 seconds ...
+ sleep 5
+ (( i++ ))
+ (( i < 120 ))
+ check_data dictionary
+ execute_sql 'select * from INCEPTOR_LIBRARY'
+ '[' 1 -ne 1 ']'
+ for hostPort in ${MYSQL_SERVER_PORT//,/ }
++ echo 127.0.0.1:3308
++ cut -d: -f1
+ host=127.0.0.1
++ echo 127.0.0.1:3308
++ cut -d: -f2
+ port=3308
+ mysql -h 127.0.0.1 -P 3308 -u inceptoruser --password=password metastore quark1 -e 'select * from INCEPTOR_LIBRARY'
ERROR 1146 (42S02) at line 1: Table 'metastore_quark1.INCEPTOR_LIBRARY' doesn't exist
+ return 1
+ echo 'Not ready, wait for 5 seconds ...'
Not ready, wait for 5 seconds ...
+ sleep 5
+ (( i++ ))
+ (( i < 120 ))
+ check_data dictionary
+ execute_sql 'select * from INCEPTOR_LIBRARY'
+ '[' 1 -ne 1 ']'

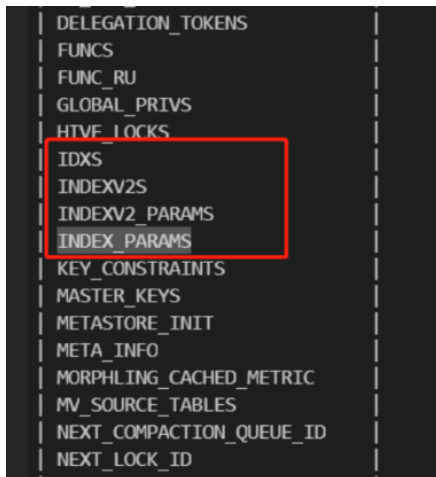
```

此时查看返回信息，信息显示缺失 `INCEPTOR_LIBRARY` 表。因此需要修复该问题。

5. 确认报错原因

连接数据库，指令：

```
mysql -h 127.0.0.1 -P 3308 -u inceptoruser --password=password metastore_quark1 show tables;
```



发现上述缺失的表确实不存在，因此需要进行修复。

6. 问题修复

执行以下SQL命令修复表。

```
create table if not exists INCEPTOR_LIBRARY(library_name varchar(200) PRIMARY KEY,
library_version varchar(100));
insert into INCEPTOR_LIBRARY values('data-dictionary',1.0);
```

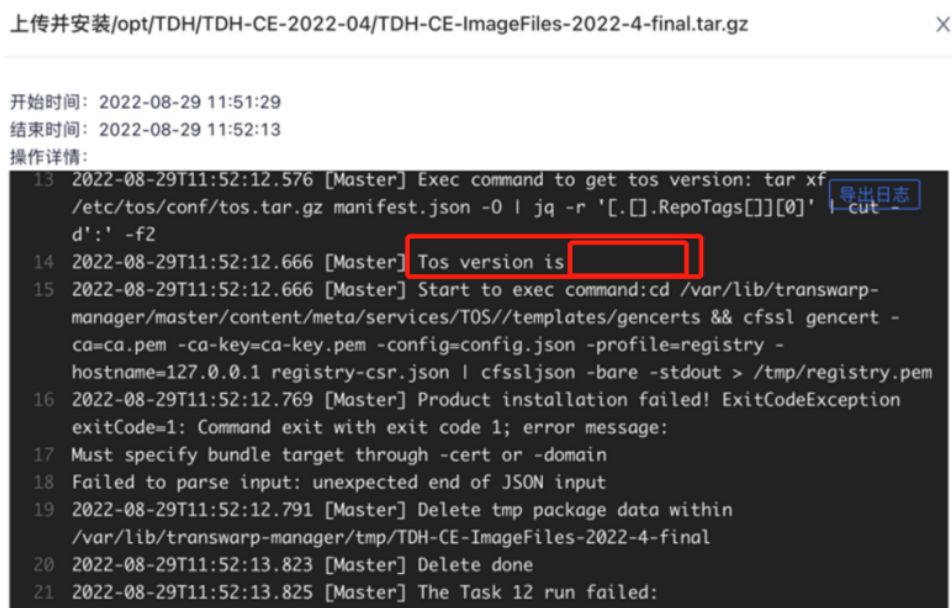
7. 重试

执行后回到Manager界面进行重试，重试后即可启动Quark。问题解决。

7.2. 上传产品包失败

• 根因1:

通过报错错误信息中TOS version is显示为空，可以判断出用户还没添加节点就上传产品包了，如下图所示。因此添加完节点再上传就可以了



• 根因2:

目录磁盘性能较差，上传超时

如果超过20分钟，则会显示失败，解决方法:

- a. 解决磁盘io弱的问题;
- b. 或者联系工作人员单独出私包，将时间修改为60分钟;



• 根因3:

根目录磁盘空间不足，解压失败

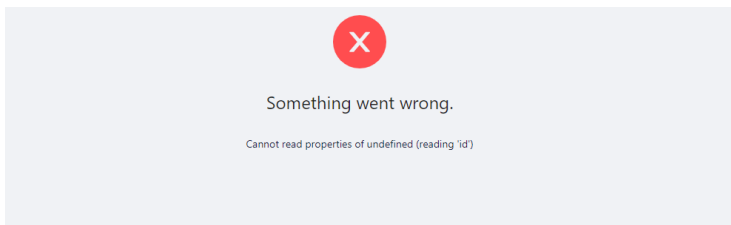
• 根因4:

产品包并非标准格式，可能是传输错误，或者是上传了其他的zip包没有解压

• 根因5:

集群registry (TOS的角色) 异常

7.3. 许可证页面打不开、上传产品包失败



操作名称	状态	提交时间	开始时间
安装产品包/home/sinitek/HYPERBASE-CE-ImageFiles-2022-10-final.tar.gz	FAILED	2023-05-24 13:40:34	2023-05-24 13:4
安装产品包/home/sinitek/TDH-CE-ImageFiles-2023-3-final.tar.gz	FAILED	2023-05-24 13:38:45	2023-05-24 13:3

上传并安装/home/sinitek/TDH-CE-ImageFiles-2023-3-final.tar.gz

X

开始时间: 2023-05-24 14:04:22

结束时间: 2023-05-24 14:05:11

操作详情:

```

137         at
      io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask
      (TaskDriver.java:224)
138     ... 10 more
139 Caused by: io.transwarp.manager.common.shell.BaseExecutor$ExitCodeException:
      ExitCodeException exitCode=125: Command exit with exit code 125; error message:
140 Unable to find image 'transwarp/registry:tos-2.1' locally
141 docker: Error response from daemon: pull access denied for transwarp/registry,
      repository does not exist or may require 'docker login'.
142 See 'docker run --help'.
143         at
      io.transwarp.manager.common.shell.BaseExecutor.executeOrFailure (BaseExecutor.jav
      a:59)
144         at
      io.transwarp.manager.master.operation.execution.localrunner.ProductInstallTaskLo
      ...
  
```

【解决方法】

TOS为集群基础服务，首次配置时，在配置服务点需要先配置TOS服务。该服务会配置镜像地址等等，上面问题均会解决

7.4. 安装过程中，配置Quark失败，报错步骤为“生成Collector配置，配置Quark Metastore”

状态	描述	开始时间	结束时间	操作
SUCCESSFUL	生成依赖的服务的配置	2023-05-10 11:43:53	2023-05-10 11:43:53	查看日志
FAILED	生成 Collector 配置	2023-05-10 11:43:53	2023-05-10 11:43:53	重试 查看日志
SUCCESSFUL	创建数据库	2023-05-10 11:43:53	2023-05-10 11:43:53	查看日志
FAILED	配置Quark Metastore(tdh1)	2023-05-10 11:43:53	2023-05-10 11:43:53	重试 查看日志
FAILED	配置Quark Metastore(tdh3)	2023-05-10 11:43:53	2023-05-10 11:43:53	重试 查看日志
WAITING	配置Quark Server(tdh1)			查看日志
WAITING	配置Nucleon Executor(tdh1)			查看日志
WAITING	配置Nucleon Executor(tdh3)			查看日志
WAITING	配置Nucleon Executor(tdh2)			查看日志
WAITING	通知 Aquila 重载配置文件			查看日志
> 初始化Quark1 0%				
> 启动Quark1 0%				
> 初始化 QUARK 预定义监控信息 0%				

【核心原因】

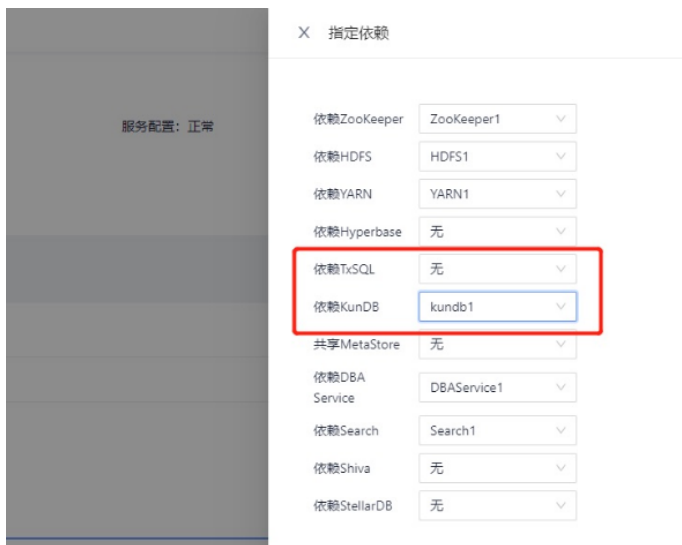
有可能是没装好kundb，所以quark没有依赖上



注意是kundb不是globalkundb哦，kundb与globalkundb服务区别：
<https://community.transwarp.cn/article/1024>

【解决方法】

- 安装kundb, 更新quark依赖, 下推配置, 重启quark。

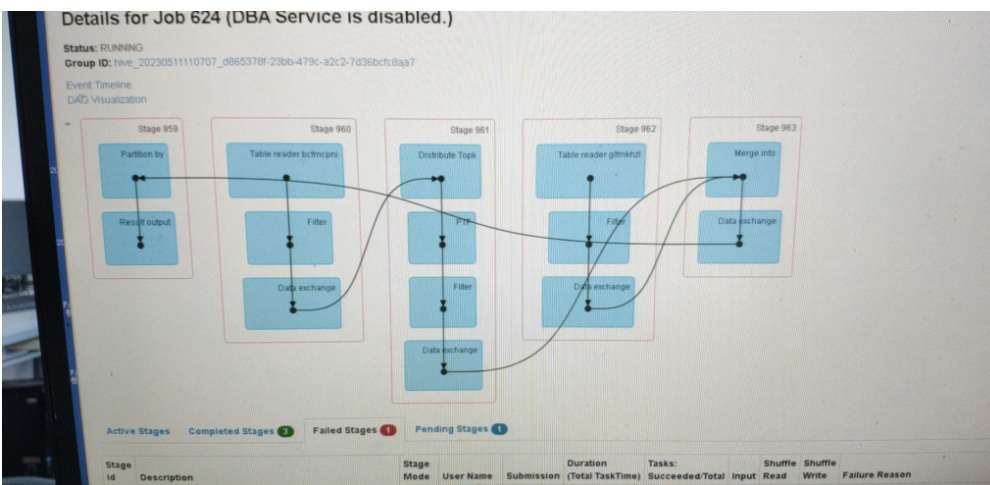


7.5. 安装Transwarp Inceptor的时候, 配置Quark Metastore(tdh1)的job报错

```

1 2023-05-10T11:43:53.197 [Master] ===== Task 264 start to run. =====
2 2023-05-10T11:43:53.197 [Master] Starting task local part ...
3 2023-05-10T11:43:53.197 [Master] Start to generate service OPS collector configs
4 2023-05-10T11:43:53.264 [Master] Waiting for association between Manager and Agent ...
5 2023-05-10T11:43:53.265 [Master] Agent tdh1 associated
6 2023-05-10T11:43:53.265 [Master] Rendering content of /etc/tdh-exporter/quark1/quark1-quark-compaction-txsq1.yml ...
7 2023-05-10T11:43:53.276 [Master] The Task 264 run failed: java.lang.RuntimeException: org.jooq.Lambda.UncheckedException: freemarker.core.InvalidReferenceException: The following has
evaluated to null or missing:
8 --> service.roles.INCEPTOR_MYSQL [in nameless template at line 28, column 3]
9
10 ----
11 Tip: It's the step after the last dot that caused this error, not those before it.
12 ----
13 Tip: If the failing expression is known to legally refer to something that's sometimes null or missing, either specify a default value like myOptionalVar!myDefault, or use <#if
myOptionalVar??>when-present{#else}>when-missing{/#if}. (These only cover the last step of the expression; to cover the whole expression, use parenthesis: (myOptionalVar.foo)
!myDefault, (myOptionalVar.foo)?)
14 ----
15
16 ----
17 FTL stack trace ("-" means nesting-related):
18 - Failed at: #assign mysqlHostPorts = [service.rol... [in nameless template at line 28, column 3]
19 ----
20 at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask(TaskDriver.java:226)
21 at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$FastClassBySpringCGLIB$$e353857f.invoke(<generated>)
22 at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
23 at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:783)
24 at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
25 at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:753)
26 at org.springframework.aop.interceptor.AsyncExecutionInterceptor.Lambda$invoke$0(AsyncExecutionInterceptor.java:115)
27 at java.util.concurrent.FutureTask.run(FutureTask.java:266)
28 at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
29 at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
30 at java.lang.Thread.run(Thread.java:748)
31 Caused by: org.jooq.Lambda.UncheckedException: freemarker.core.InvalidReferenceException: The following has evaluated to null or missing:
32 --> service.roles.INCEPTOR_MYSQL [in nameless template at line 28, column 3]
33
34 ----
35 Tip: It's the step after the last dot that caused this error, not those before it.
36 ----
37 Tip: If the failing expression is known to legally refer to something that's sometimes null or missing, either specify a default value like myOptionalVar!myDefault, or use <#if
myOptionalVar??>when-present{#else}>when-missing{/#if}. (These only cover the last step of the expression; to cover the whole expression, use parenthesis: (myOptionalVar.foo)
!myDefault, (myOptionalVar.foo)?)
38 ----
39
40 ----
41 FTL stack trace ("-" means nesting-related):
42 - Failed at: #assign mysqlHostPorts = [service.rol... [in nameless template at line 28, column 3]
43 ----
44 at org.jooq.Lambda.Unchecked.Lambda$static$0(Unchecked.java:61)

```



Duration (Total TaskTime)	Tasks: Succeeded/Total	Shuffle Input	Shuffle Read	Shuffle Write	Failure Reason
6.3 min (28 s)	45/61 (49 failed)	60.1 MB	56.5 MB		Job aborted due to stage failure: Task 8 in stage 998.0 failed 4 times, most recent failure: Lost task 8.3 in stage 998.0 (TID 22678, 10.224.20.0): FetchFailed(null, shuffleId=20974, mapId=-1, reduceId=22) Driver stacktrace:
72 ms (0 ms)	0/5				Job aborted due to stage failure: Task 8 in stage 998.0 failed 4 times, most recent failure: Lost task 8.3 in stage 998.0 (TID 22678, 10.224.20.0): FetchFailed(null, shuffleId=20974, mapId=-1, reduceId=22) Driver stacktrace:

【核心原因】

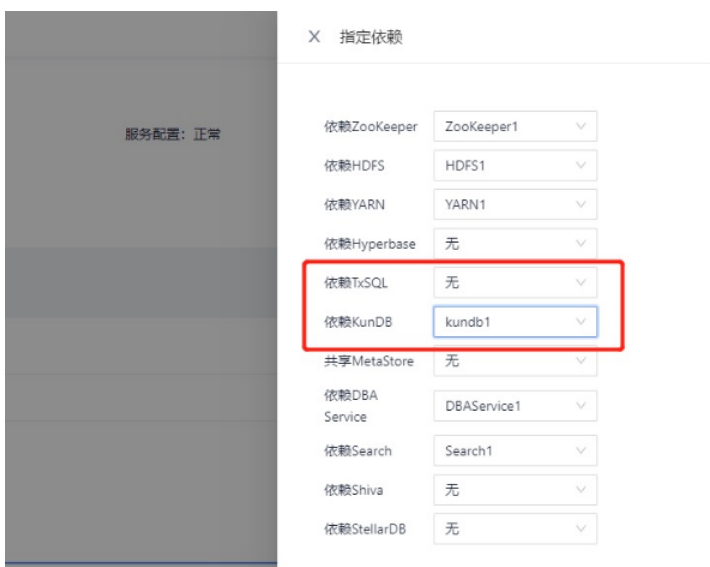
- Quark没有依赖KunDB

【解决方法】

- 更新quark依赖，下推配置，重启quark。



此处需要注意的是kundb不是globalkundb哦， kundb与globalkundb服务区别：
<https://community.transwarp.cn/article/1024>



7.6. 安装Guardian服务报错：“/api/market/install/advices: 500 Internal Server Error”

【核心原因】

Guardian是单独的产品服务，有独立的产品包，出现该报错的原因是因为没有上传对应的产品包，产品包

同样在官网下载弹窗中获取

【解决方法】

下载上传后进行添加

7.7. 安装Guardian服务报错：“/api/services/19:500 Internal Server Error”

【问题描述】

最开始因为没有上传产品包报错了，后来上传后在配置服务的步骤报错了。



【核心原因】

操作不当

【解决方法】

删除，重新安装

7.8. 安装guardian的时候报错显示“error code 48000…unauthorized”

【报错信息】

```

同: 2022-10-26 19:19:23
同: 2022-10-26 19:19:23
情:
2022-10-26T06:17:45.349 [Master] ===== Task 1331 start to run. =====
2022-10-26T06:17:45.350 [Master] Starting task local part ...
2022-10-26T06:17:45.358 [Master] Start add federation user for manager
2022-10-26T06:17:45.602 [Master] User manager already exists, do nothing
2022-10-26T06:17:45.602 [Master] Add federation user for manager done
2022-10-26T06:17:45.605 [Master] Start register tenant TDH
2022-10-26T06:17:45.656 [Master] The Task 1331 run failed:
java.lang.RuntimeException:
io.transwarp.guardian.federation.client.FederationClientException:
{"error": "Unauthorized", "error_code": 48000, "error_description": "Unauthorized"}
    at
    io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask
(TaskDriver.java:226)
    at
    io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$FastCl
    ...
    
```

【解决方法】

重新开启一个页面进入Guardian服务，角色栏中点击‘Guardian Federation Service’右侧的查看中的链接，进入登录页面，选择用“平台用户登录”，登录后进到用户管理页面，把“manager”这个用户删了，然后返回报错的页面点击“重试”即可。

<input type="checkbox"/> 角色名称	节点名称	机柜名称	健康状况	服务链接	操作
<input type="checkbox"/> Guardian Federation Service (Guardian, idc2611)	idc2611	/default-rack	● Running	查看	启动 停止 删除 ...
<input type="checkbox"/> Guardian Server (Guardian, idc2611)	idc2611	/defau		https://172.18.126.11:8383/federation-server?tenantName=TDH2611	



用户	描述	角色	创建时间	操作
admin admin@Federation	Federation administrative user	系统管理员 系统租户管理员	2021/11/6 11:49:41	编辑 删除
manager manager@transwarp.io		系统管理员	2021/11/6 11:53:18	编辑 删除

7.9. 第二次安装Guardian时卡住了，报错显示“error code 403... Access is denied”

【报错信息】

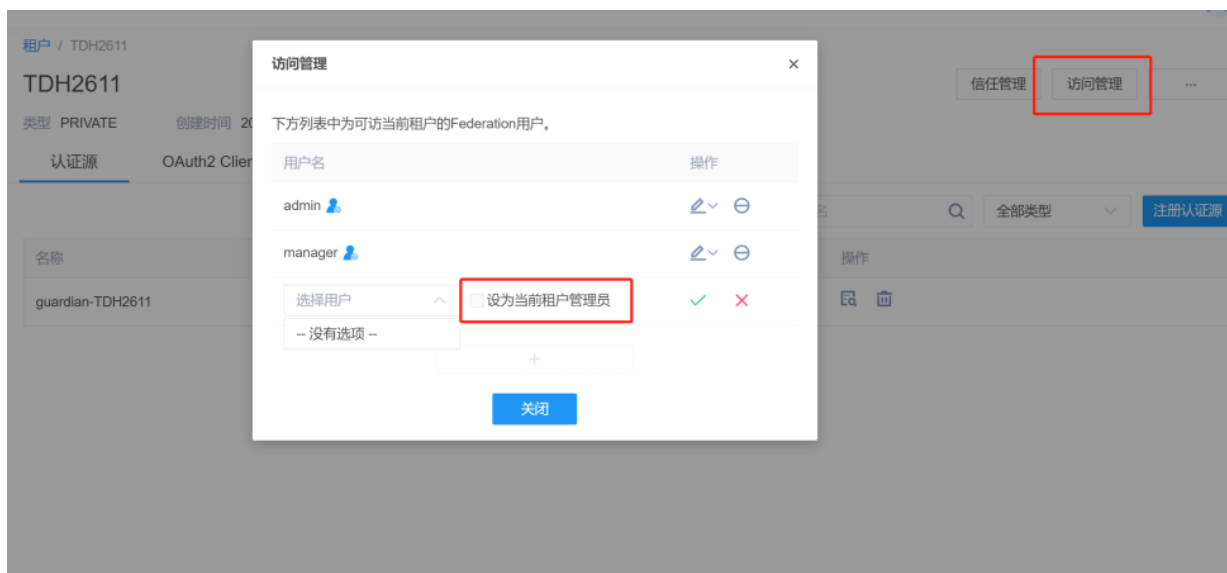

```

时间: 2022-10-28 10:05:15
时间: 2022-10-28 10:05:15
详情:
0 2022-10-27T22:05:15.214 [Master] ===== Task 3149 start to run. =====
1 2022-10-27T22:05:15.214 [Master] Starting task local part ...
2 2022-10-27T22:05:15.218 [Master] Start add federation user for manager
3 2022-10-27T22:05:15.468 [Master] Add federation user for manager done
4 2022-10-27T22:05:15.471 [Master] Start register tenant TDH
5 2022-10-27T22:05:15.548 [Master] Tenant Tenant{name='TDH', type=PRIVATE,
attributes={}, createTime=1666922715471, description='null'} already exists, do
noting
6 2022-10-27T22:05:15.548 [Master] Register tenant TDH done
7 2022-10-27T22:05:15.548 [Master] Start set tenant access for admin user
8 2022-10-27T22:05:15.618 [Master] The Task 3149 run failed:
java.lang.RuntimeException:
io.transwarp.guardian.federation.client.FederationClientException:
{"error": "Forbidden", "error_code": 403, "error_description": "Access is denied"}
at
io.transwarp.manager.master.operation.execution.TaskDriverTaskExecutor.runTask

```

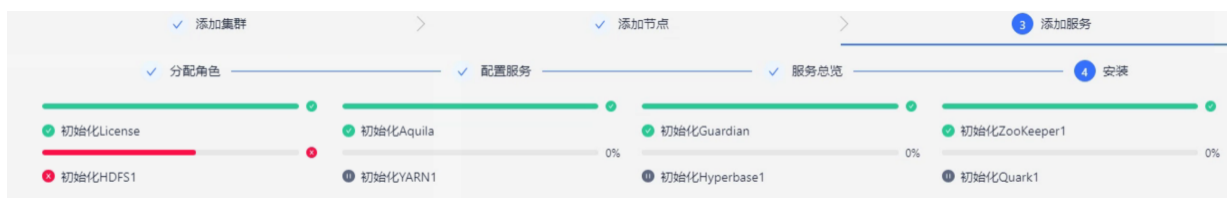
【解决方法】

重新开启一个页面访问Guardian Server进到对应的TDH的租户里面，点击访问管理，点击‘+’把manager用户加成这个租户的管理员，然后返回报错的页面点击“重试”即可。



7.10. 社区开发版安装过程中，初始化HDFS失败

【报错信息】



```

2023-06-10T15:09:56.046073 [Master] Start executing [chmod 755 /var/lib/transwarp-manager/master/content/resources/nodes/localhost/@etc@tdh-exporter@hdfs1@HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml]
2023-06-10T15:09:56.057937 [Master] Execute success.
2023-06-10T15:09:56.058090 [Master] Start copy from [localhost] /var/lib/transwarp-manager/master/content/resources/nodes/localhost/@etc@tdh-exporter@hdfs1@HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml to [localhost] /var/lib/transwarp-manager/agent/resource-tmp/@etc@tdh-exporter@hdfs1@HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml
2023-06-10T15:09:56.064942 [Master] Copy success.
2023-06-10T15:09:56.084402 [Master] Task local part ended.
2023-06-10T15:09:56.084609 [Master] Starting task remote part ...
2023-06-10T15:09:56.088155 [Agent] Execute command: systemctl start transwarp-hdfs-datanode@hdfs1
2023-06-10T15:09:56.299771 [Agent] Execute command: umask 0022 && mkdir -p $(dirname "/etc/tdh-exporter/hdfs1/HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml") && mv -f "/var/lib/transwarp-manager/agent/resource-tmp/@etc@tdh-exporter@hdfs1@HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml" "/etc/tdh-exporter/hdfs1/HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml" && chmod 755 "/etc/tdh-exporter/hdfs1/HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml" && chown hdfs "/etc/tdh-exporter/hdfs1/HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml" && chown :hdfs "/etc/tdh-exporter/hdfs1/HDFS_DATANODE-14-hdfs-datanode-HttpJson.yml"
2023-06-10T15:09:56.362222 [Master] Waiting Data Node (HDFS1,localhost) to become Healthy within 600 s
2023-06-10T15:12:55.889723 [Master] Task 67 timed out after 180000ms.
2023-06-10T15:12:55.891411 [Master] The Task 67 run failed: java.util.concurrent.CancellationException
at java.base/java.util.concurrent.FutureTask.report(FutureTask.java:122)
at java.base/java.util.concurrent.FutureTask.get(FutureTask.java:191)
at org.springframework.util.concurrent.ListenableFutureTask.done(ListenableFutureTask.java:83)
at java.base/java.util.concurrent.FutureTask.finishCompletion(FutureTask.java:381)
at java.base/java.util.concurrent.FutureTask.cancel(FutureTask.java:179)
at io.transwarp.manager.master.operation.execution.TaskDriver.lambda$submitTask$0(TaskDriver.java:79)
at java.base/java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:515)
at java.base/java.util.concurrent.FutureTask.run(FutureTask.java:264)

```

【核心原因】

未更改主机名

【解决方法】

请在服务器端使用 `hostnamectl set-hostname [hostname]`。主机名注意需符合DNS-1123规范，由数字、小写字母或“-”组成，不能包含大写字母，长度小于63

7.11. Dashboard安装完成后状态显示不可用

【报错信息】



【解决方法】

- 先重启

Dashboard

重启Dashboard (2/2)

状态	描述	开始时间	结束时间	操作
成功	为Dashboard(zpybdp02)移除TOS标签	2022-06-29 16:15:04	2022-06-29 16:15:05	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:05	2022-06-29 16:15:05	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:05	2022-06-29 16:15:05	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:05	2022-06-29 16:15:05	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:05	2022-06-29 16:15:06	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:06	2022-06-29 16:15:06	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:06	2022-06-29 16:15:06	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:06	2022-06-29 16:15:07	查看日志
成功	通过TOS停止Dashboard	2022-06-29 16:15:07	2022-06-29 16:15:07	查看日志
成功	为Dashboard(zpybdp02)添加TOS标签	2022-06-29 16:15:07	2022-06-29 16:15:07	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:07	2022-06-29 16:15:08	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:08	2022-06-29 16:15:08	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:08	2022-06-29 16:15:09	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:09	2022-06-29 16:15:10	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:10	2022-06-29 16:15:10	查看日志
成功	通过TOS启动Dashboard	2022-06-29 16:15:10	2022-06-29 16:15:11	查看日志
失败	通过TOS启动Dashboard	2022-06-29 16:15:11	2022-06-29 16:15:11	重试 查看日志

- 如果重启后依然显示不可用的话，修改/var/lib/transwarp-manager/master/content/meta/services/DASHBOARD/tos-2.1/templates/下面的dashboard-deployment.yaml和dashboard-service.yaml文件：

a. dashboard-deployment.yaml文件这么改

```

- key: kubernetes-dashboard
  operator: Exists
containers:
- name: kubernetes-dashboard
  image: transwarp/dashboard:tos-2.0
  imagePullPolicy: Always
  args:
    # Uncomment the following line to manually specify Kubernetes API server host and port:
    # - --api-server=http://localhost:8080
- key: kubernetes-dashboard
  operator: Exists
containers:
- name: kubernetes-dashboard
  image: transwarp/dashboard:tos-2.1
  imagePullPolicy: Always
  args:
    # Uncomment the following line to manually specify Kubernetes API server host and port:
    # - --api-server=http://localhost:8080

```

b. dashboard-service.yaml文件这么改

```

kind: Service
apiVersion: v1
metadata:
  labels:
    name: kubernetes-dashboard
  name: kubernetes-dashboard
  namespace: kube-system
spec:
  type: NodePort
  ports:
    - port: ${service['tos.master.dashboard.port']}
      targetPort: ${target['tos.master.dashboard.port']}
      nodePort: ${service['tos.master.dashboard.port']}
  selector:
    name: kubernetes-dashboard
kind: Service
apiVersion: v1
metadata:
  labels:
    name: kubernetes-dashboard
  name: kubernetes-dashboard
  namespace: kube-system
spec:
  type: NodePort
  ports:
    - port: ${service['dashboard.ssl.container.port']}
      targetPort: ${service['dashboard.ssl.container.port']}
      nodePort: ${service['tos.master.dashboard.port']}
  selector:
    name: kubernetes-dashboard

```

然后再重启Manager，重启后选择右上角配置服务，配置完成后再Dashboard配置服务，之后重启Dashboard就好了

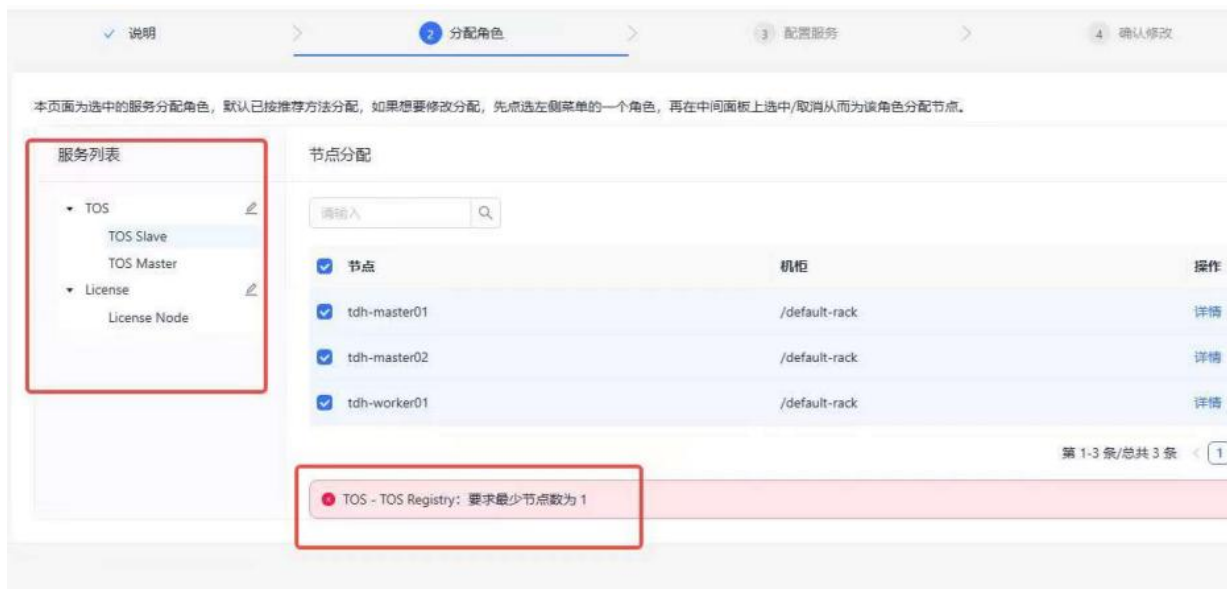


注：dashboard配置服务在服务右上角的菜单里

7.12. TOS服务显示待安装状态，安装时报错无法添加register角色

【问题描述】

TDH第一次安装，在安装tos的时候显示 待安装状态，然后添加角色显示tos-tos registry：要求最少节点数为1，但是加不上register角色，这样要怎么解决？



【解决方法】

TOS服务未安装成功，可以在TOS服务页面删除服务后，重新安装服务。安装流程：集群—添加服务—其他—TOS。

8. 使用过程中可能遇到的问题

8.1. 某个服务突然挂掉了的常规排查步骤

1. 查看基础服务（LicenseServer/TOS/Aquila）是否健康，因为产品服务是运行在基础服务之上，如果基础服务不健康是会导致产品服务无法运行，因此优先解决基础服务的健康问题；
2. 如果基础服务健康，但是单独产品服务有问题，则进入服务详情页看具体是哪个角色有问题，然后在/var/log目录下看对应的日志是否有报错，比如端口冲突等信息；
3. 如果后台日志里看不出来问题，再看manager日志（/var/log/transwarp-manager），前台日志（任务日志及具体服务的健康状态旁的详情）；
4. 如果日志显示是pod出问题了，那么就可以执行kubect1 get pod |grep <服务名字>, kubect1 logs <不running的podid>, 来进一步查看pod的日志。

如果依然无法判断出问题，可以访问星环开发者社区联系社区版工作人员，并提供相关的报错日志



8.2. 某个服务启动失败的常规排查步骤

当某个服务出现问题的时候可以参考下方思路进行问题定位：

1. 重试当前任务，看看是否能够启动成功（这个是因为可能是因为磁盘io或其他原因导致超时报错，不一定是服务本身出现问题）
2. 返回服务主界面，选择重启服务，看看是否能够重启成功
3. 通过 `systemctl list-units | grep <组件名, 如quark1>`，查找服务角色对应的service名称
4. 通过 `systemctl status <service_name>` 查看服务角色状态
5. 查看服务日志，看是否有明显报错
6. 若无日志生成，通过 `journalctl -u <service_name>` 查看服务启动日志，看是否有明显报错

8.2.1. 示例1.启动TDH服务失败如何定位问题

【问题描述】

当在8180页面上发现服务不健康（红色或者黄色）或者未启动（灰色）时，常见的办法是去启动（或者重

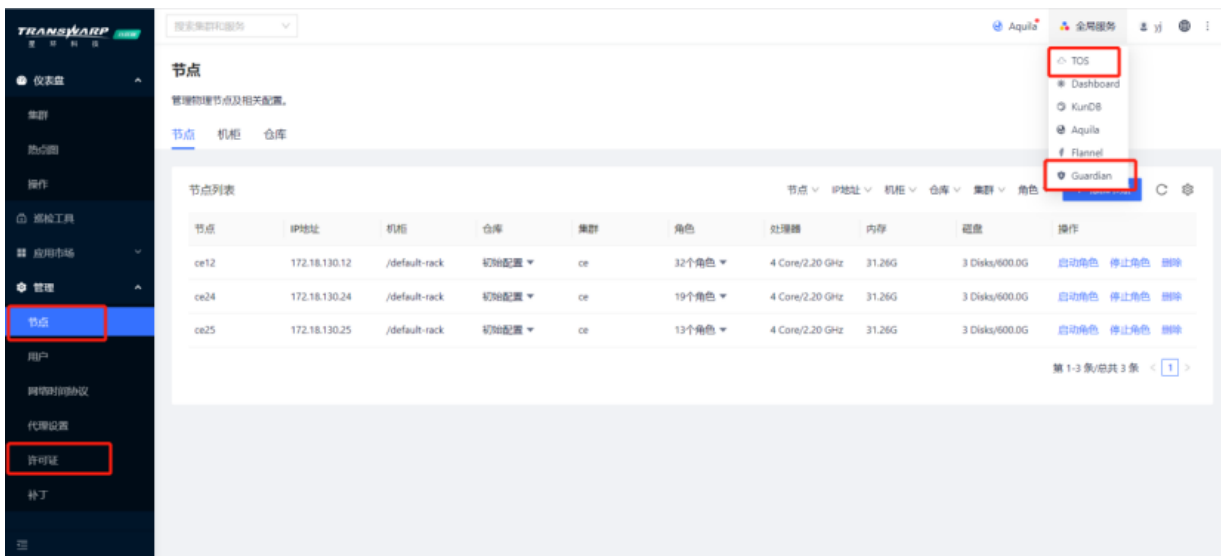
启) 问题服务，服务启动失败时需要检查两个方面：

1. 容器相关的服务是否正常
2. 服务本身的进程启动情况

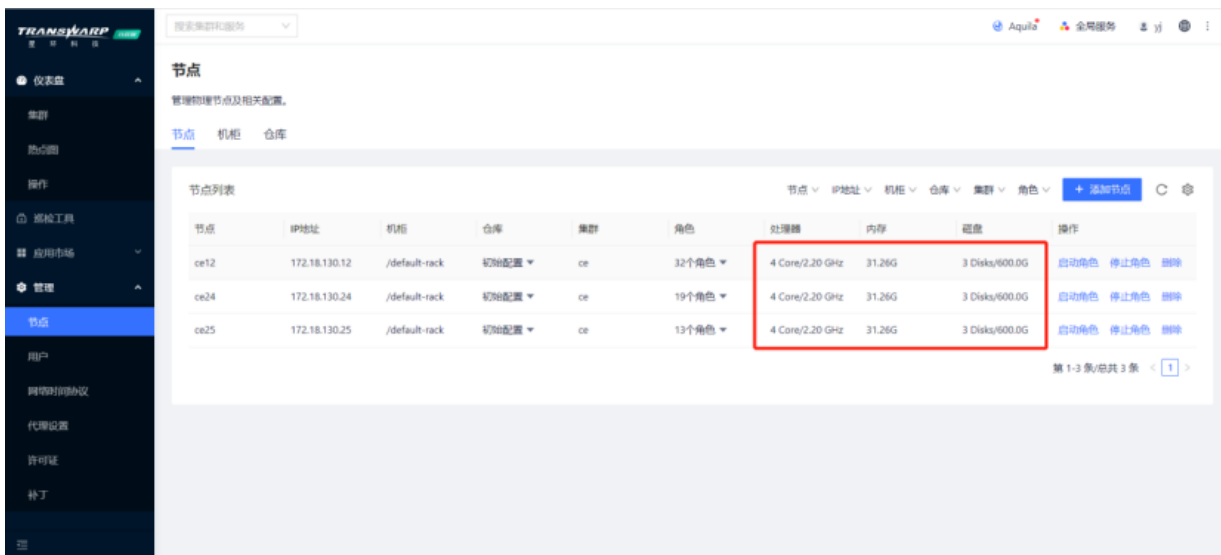
【具体步骤】

步骤一 检查集群基础的服务是否健康

TDH服务的健康依赖于许多的基础组件，当发现启动服务失败时候，请首先检查这些基础服务的健康状态，一般包括transwarp-manager-agent/tos/guardian/许可证等服务



如果有不健康的可以通过页面来启动它，需要注意的是节点页面上必须要能显示硬件信息才证明是健康的，如下图所示



如果没有显示硬件信息的话则需要ssh登录到对应的服务器上，手动启动一下agent进程：

```
systemctl restart transwarp-manager-agent
```

步骤二 查看页面上报错的原因

任何一个操作失败的时候，在页面上都会给出对应的输出，所以排查的第一步就是点击“查看”这些输出中的报错：

配置节点tdh-02的NTP服务

x

开始时间:2022/06/24 15:22:14

完成时间:2022/06/24 15:22:44

状态:FAILED

详情:

```

Fri Jun 24 15:22:14 GMT+08:00 2022 [Manager] Starting task local part ...
ntp.synchronization.strategy is set as external_and_internal
Fri Jun 24 15:22:14 GMT+08:00 2022 [Agent] execute command: systemctl stop ntpd
akka.pattern.AskTimeoutException: Ask timed out on [ActorSelection[Anchor(akka.tcp://agent@tdh-02:10208/),
Path(/user/agentDispatcher)]] after [30000 ms]
    at akka.pattern.PromiseActorRef$$anonfun$1.apply$mcV$sp(AskSupport.scala:334)
    at akka.actor.Scheduler$$anon$7.run(Scheduler.scala:117)
    at
scala.concurrent.Future$InternalCallbackExecutor$.scala$concurrent$Future$InternalCallbackExecutor$$unb
atchedExecute(Future.scala:694)
    at scala.concurrent.Future$InternalCallbackExecutor$.execute(Future.scala:691)
    at akka.actor.LightArrayRevolverScheduler$TaskHolder.executeTask(Scheduler.scala:467)
    at akka.actor.LightArrayRevolverScheduler$$anon$8.executeBucket$1(Scheduler.scala:419)
    at akka.actor.LightArrayRevolverScheduler$$anon$8.nextTick(Scheduler.scala:423)
    at akka.actor.LightArrayRevolverScheduler$$anon$8.run(Scheduler.scala:375)

```

比如上图中的输出就是在提示连接对应节点的10208端口失败了，接下来只要找到这个端口对应的进程把它启动即可。

当然现实往往不会很完美，当遇到的是启动服务失败的情况时，页面上的输出往往不会如此直接指向错误的根源：

通过TOS启动server

x

开始时间:2022/06/03 12:03:22

完成时间:2022/06/03 12:06:24

状态:FAILED

详情:

```

server.yaml
start to generate mysql-server.yaml on [Manager]...
start handle local part, dataModel see transwarp-manager.log
generated mysql-server.yaml on [Manager]
start to create role(s) on [Manager] using kubectl --server=https://127.0.0.1:6443 --certificate-authority=/srv
/kubernetes/ca.pem --client-certificate=/srv/kubernetes/admin.pem --client-key=/srv/kubernetes/admin-key.pem
scale --replicas=1 -f /var/lib/transwarp-manager/master/content/resources/services/mysql1/mysql-server.yaml...
role(s) successfully created on [Manager]
Fri Jun 03 12:03:24 GMT+08:00 2022 [Manager] Task local part ended
Waiting MYSQL_SERVER in MySQL1 to become Healthy within 600 seconds ...
Latest health check result of roles:
DAEMON_CHECK DOWN at Fri Jun 03 12:06:24 GMT+08:00 2022
  MYSQL_SERVER on tdh-01 has Pod mysql-server-mysql1-56458fc5c7-jzcw8 with status CrashLoopBackOff
io.transwarp.manager.master.manager.operation.TaskLocalRunner$DownException: MYSQL_SERVER in
MySQL1 didn't become Healthy within 600 seconds

```

步骤三 查看服务的后台pod状态及日志

自TDH发布5.x版本以后，所有服务都是容器化的，所以当服务启动失败时候，先检查它的容器pod状态是否正常，比如如果是inceptor服务启动失败，那么可以在后台执行：


```
kubectl get pod -owide | grep quark1
```

```
[root@ce24 quark1]# kubectl get pod -owide | grep quark1
nucleon-executor1-quark1-66df7967cb-rjs68      1/1      Running    0          44d      172.18.130.25   ce25      <none>
<none>
nucleon-executor1-quark1-66df7967cb-vkc4r      1/1      Running    0          44d      172.18.130.24   ce24      <none>
<none>
nucleon-executor1-quark1-66df7967cb-xrgsj      1/1      Running    0          44d      172.18.130.12   ce12      <none>
<none>
quark-metastore-quark1-7d86cd8b8c-6w7cf       1/1      Running    0          44d      172.18.130.12   ce12      <none>
quark-metastore-quark1-7d86cd8b8c-cwg24       1/1      Running    0          44d      172.18.130.24   ce24      <none>
quark-server-quark1-99866bbcc-t6j7j           1/1      Running    0          44d      172.18.130.24   ce24      <none>
<none>
```

健康

状态

结果中的状态必须是running、健康探针必须是1/1时这个pod才是健康的，否则就需要去查看这个pod的日志，命令一般是两条：

```
kubectl logs <podid>
kubectl describe pod <podid>
```

```
[root@ce24 quark1]# kubectl get pod |grep quark1
nucleon-executor1-quark1-66df7967cb-rjs68      1/1      Running    0          44d
nucleon-executor1-quark1-66df7967cb-vkc4r      1/1      Running    0          44d
nucleon-executor1-quark1-66df7967cb-xrgsj      1/1      Running    0          44d
quark-metastore-quark1-7d86cd8b8c-6w7cf       1/1      Running    0          44d
quark-metastore-quark1-7d86cd8b8c-cwg24       1/1      Running    0          44d
quark-server-quark1-99866bbcc-t6j7j           1/1      Running    0          44d
[root@ce24 quark1]# kubectl logs quark-metastore-quark1-7d86cd8b8c-6w7cf
+ export TDH_SCRIPT_BIN=/bin
+ TDH_SCRIPT_DIR=/bin
+ source /usr/lib/transwarp/scripts/repeat_until_ready.sh
umount /etc/hosts
+ '[' x '!=' xOVERLAY ']'
+ echo 'umount /etc/hosts'
+ umount /etc/hosts
rm -rf /etc/hosts
+ echo 'rm -rf /etc/hosts'
+ rm -rf /etc/hosts
+ echo 'ln -sf /etc/transwarp/conf/hosts /etc/'
+ ln -s /etc/transwarp/conf/hosts /etc/
ln -sf /etc/transwarp/conf/hosts /etc/
+ DEBUG=0
```

```
2022-12-27 17:30:08,347: ...useoptimizehiveconf: false
log4j:WARN No such property [maxBackupIndex] in org.apache.hadoop.hive.ql.log.InceptorRollingFileAppender.
log4j:WARN No such property [maxFileSize] in org.apache.hadoop.hive.ql.log.InceptorRollingFileAppender.
log4j:WARN No such property [maxBackupIndex] in org.apache.hadoop.hive.ql.log.InceptorRollingFileAppender.
log4j:WARN No such property [maxFileSize] in org.apache.hadoop.hive.ql.log.InceptorRollingFileAppender.
Loading class 'com.mysql.jdbc.Driver'. This is deprecated. The new driver class is 'com.mysql.cj.jdbc.Driver'. The driver is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.
[root@ce24 quark1]# kubectl describe pod quark-metastore-quark1-7d86cd8b8c-6w7cf
Name:          quark-metastore-quark1-7d86cd8b8c-6w7cf
Namespace:    default
Priority:      0
Priority Class Name:  low-priority
Node:         ce12/172.18.130.12
Start Time:   Tue, 27 Dec 2022 17:29:35 +0800
Labels:       name=quark-metastore-quark1
              pod-template-hash=7d86cd8b8c
              podConflictName=quark-metastore-quark1
Annotations:  serviceInstanceName: Quark1
Status:       Running
IP:           172.18.130.12
IPs:
IP:           172.18.130.12
Controlled By:  ReplicaSet/quark-metastore-quark1-7d86cd8b8c
Containers:
  quark-metastore-quark1:
    Container ID:  docker://64377fa8e0d2c09ef7dba7efdce9eaea23b65ae59eb62231e9b8cde83cf8037b
    Image:         transwarp/quark:quark-8.23-ce-final
    Image ID:     docker-pullable://ce12:5000/transwarp/quark@sha256:a533f6cc8478dd13223b4f6d78a313c96babebda6affdbb100ebe8020e08b9db
    Port:         <none>
```

两条命令输出的日志侧重点并不相同，根据结果中的错误信息往往能够对服务启动的原因可以有初步的判断。

步骤四 查看服务的日志

如果从pod的日志里还是无法找到错误，可以直接去检查服务的日志。分为以下两种情况：

- 第一种情况：Pod没有起来，显示0/1，那么首先应该根据上一步的输出结果查到问题pod所在的节点

```
[root@ce24 quark1]# kubectl get pod -owide | grep quark
nucleon-executor1-quark1-66df7967cb-rjs68          1/1    Running    0         44d    172.18.130.25    ce25    <none>
<none>
nucleon-executor1-quark1-66df7967cb-vkc4r          1/1    Running    0         44d    172.18.130.24    ce24    <none>
<none>
nucleon-executor1-quark1-66df7967cb-xrgsj          0/1    Running    0         44d    172.18.130.12    ce12    <none>
<none>
quark-metastore-quark1-7d86cd8b8c-6w7cf           1/1    Running    0         44d    172.18.130.12    ce12    <none>
<none>
quark-metastore-quark1-7d86cd8b8c-cwg24           1/1    Running    0         44d    172.18.130.24    ce24    <none>
<none>
quark-server-quark1-99866bbcc-t6j7j               1/1    Running    0         44d    172.18.130.24    ce24    <none>
<none>
```

登录到对应的服务器之后，可以到这个路径下查找日志：/var/log/服务名。例如quark服务的路径一般都是/var/log/quark1/（服务名称最后数字是安装的时候决定的）

- 第二种情况：如果pod起来了显示的是1/1那么则没有节点的限制（进入pod内就会切换到server所在节点），同样是到对应路径下查找日志：/var/log/服务名。

一般来说，该目录下往往会有多个角色的日志，具体需要查看哪一个文件呢？

一个角色的名字往往都是有一致性的，从pod名跟文件名就可以得出其关联，比如：

```
Node-Selectors: quark-metastore-quark1=true
Tolerations:   node.kubernetes.io/not-ready:NoExecute for 1800s
               node.kubernetes.io/unreachable:NoExecute for 1800s
Events:        <none>
[root@ce24 quark1]# kubectl get pod -owide | grep quark
nucleon-executor1-quark1-66df7967cb-rjs68          1/1    Running    0         44d    172.18.130.25    ce25    <none>
nucleon-executor1-quark1-66df7967cb-vkc4r          1/1    Running    0         44d    172.18.130.24    ce24    <none>
nucleon-executor1-quark1-66df7967cb-xrgsj          1/1    Running    0         44d    172.18.130.12    ce12    <none>
quark-metastore-quark1-7d86cd8b8c-6w7cf           1/1    Running    0         44d    172.18.130.12    ce12    <none>
quark-metastore-quark1-7d86cd8b8c-cwg24           1/1    Running    0         44d    172.18.130.24    ce24    <none>
quark-server-quark1-99866bbcc-t6j7j               0/1    CrashLoopBackoff    0         44d    172.18.130.24    ce24    <none>
[root@ce24 quark1]# cd /var/log/quark1/
0_1670984317087.jstack      driver_1670924085123.jhisto  quark-metastore.log.2023-01-09  quark-server.audit.2022-12-17  quark-server.audit.2023-01-28  quark-server.log.2023-01-04
0_1670984447408.jstack      driver_1670930226854.jhisto  quark-metastore.log.2023-01-10  quark-server.audit.2022-12-18  quark-server.audit.2023-01-29  quark-server.log.2023-01-05
13_1670919718379.jstack     jar                            quark-metastore.log.2023-01-11  quark-server.audit.2022-12-19  quark-server.audit.2023-01-30  quark-server.log.2023-01-06
15_1670919738015.jstack     log_backup.jar                quark-metastore.log.2023-01-12  quark-server.audit.2022-12-20  quark-server.audit.2023-01-31  quark-server.log.2023-01-07
15_1670920159409.jstack     log_opository_20221224_173014-706  quark-metastore.log.2023-01-13  quark-server.audit.2022-12-21  quark-server.audit.2023-02-01  quark-server.log.2023-01-08
21_1670929633853.jstack     quark-executor-0.log          quark-metastore.log.2023-01-14  quark-server.audit.2022-12-22  quark-server.audit.2023-02-02  quark-server.log.2023-01-09
3_1670905303430.jstack     quark-executor_gc.log        quark-metastore.log.2023-01-15  quark-server.audit.2022-12-23  quark-server.audit.2023-02-03  quark-server.log.2023-01-10
6_1670906246694.jstack     quark-library.log            quark-metastore.log.2023-01-16  quark-server.audit.2022-12-24  quark-server.audit.2023-02-04  quark-server.log.2023-01-11
9_1670915948330.jstack     quark-library.log.2022-08-01    quark-metastore.log.2023-01-17  quark-server.audit.2022-12-25  quark-server.audit.2023-02-05  quark-server.log.2023-01-12
driver_1670904254516.jstack quark-metastore.log.2022-12-16  quark-metastore.log.2023-01-18  quark-server.audit.2022-12-26  quark-server.audit.2023-02-06  quark-server.log.2023-01-13
driver_1670904260853.jstack quark-metastore.log           quark-metastore.log.2023-01-19  quark-server.audit.2022-12-27  quark-server.audit.2023-02-07  quark-server.log.2023-01-14
driver_1670904261864.jstack quark-metastore.log.2022-12-09  quark-metastore.log.2023-01-20  quark-server.audit.2022-12-28  quark-server.audit.2023-02-08  quark-server.log.2023-01-15
driver_1670904262863.jstack quark-metastore.log.2022-12-10  quark-metastore.log.2023-01-21  quark-server.audit.2022-12-29  quark-server.gc.log           quark-server.log.2023-01-16
driver_1670904263975.jstack quark-metastore.log.2022-12-11  quark-metastore.log.2023-01-22  quark-server.audit.2022-12-30  quark-server.gc.log.old1      quark-server.log.2023-01-17
driver_1670904264858.jstack quark-metastore.log.2022-12-12  quark-metastore.log.2023-01-23  quark-server.audit.2023-01-31  quark-server.log.old1        quark-server.log.2023-01-18
driver_1670904286855.jstack quark-metastore.log.2022-12-13  quark-metastore.log.2023-01-24  quark-server.audit.2023-01-01  quark-server.log             quark-server.log.2023-01-19
driver_1670904287894.jstack quark-metastore.log.2022-12-14  quark-metastore.log.2023-01-25  quark-server.audit.2023-01-02  quark-server.log.2022-12-09  quark-server.log.2023-01-20
driver_1670904288893.jstack quark-metastore.log.2022-12-15  quark-metastore.log.2023-01-26  quark-server.audit.2023-01-03  quark-server.log.2022-12-10  quark-server.log.2023-01-21
driver_1670904289878.jstack quark-metastore.log.2022-12-16  quark-metastore.log.2023-01-27  quark-server.audit.2023-01-04  quark-server.log.2022-12-11  quark-server.log.2023-01-22
driver_1670904290866.jstack quark-metastore.log.2022-12-17  quark-metastore.log.2023-01-28  quark-server.audit.2023-01-05  quark-server.log.2022-12-12  quark-server.log.2023-01-23
driver_1670904336007.jstack quark-metastore.log.2022-12-18  quark-metastore.log.2023-01-29  quark-server.audit.2023-01-06  quark-server.log.2022-12-13  quark-server.log.2023-01-24
driver_1670904342055.jstack quark-metastore.log.2022-12-19  quark-metastore.log.2023-01-30  quark-server.audit.2023-01-07  quark-server.log.2022-12-14  quark-server.log.2023-01-25
driver_1670904343033.jstack quark-metastore.log.2022-12-20  quark-metastore.log.2023-01-31  quark-server.audit.2023-01-08  quark-server.log.2022-12-15  quark-server.log.2023-01-26
driver_1670904344042.jstack quark-metastore.log.2022-12-21  quark-metastore.log.2023-02-01  quark-server.audit.2023-01-09  quark-server.log.2022-12-16  quark-server.log.2023-01-27
driver_1670904345043.jstack quark-metastore.log.2022-12-22  quark-metastore.log.2023-02-02  quark-server.audit.2023-01-10  quark-server.log.2022-12-17  quark-server.log.2023-01-28
driver_1670904346036.jstack quark-metastore.log.2022-12-23  quark-metastore.log.2023-02-03  quark-server.audit.2023-01-11  quark-server.log.2022-12-18  quark-server.log.2023-01-29
driver_1670904351679.jstack quark-metastore.log.2022-12-24  quark-metastore.log.2023-02-04  quark-server.audit.2023-01-12  quark-server.log.2022-12-19  quark-server.log.2023-01-30
driver_1670904374697.jstack quark-metastore.log.2022-12-25  quark-metastore.log.2023-02-05  quark-server.audit.2023-01-13  quark-server.log.2022-12-20  quark-server.log.2023-01-31
driver_1670905645332.jstack quark-metastore.log.2022-12-26  quark-metastore.log.2023-02-06  quark-server.audit.2023-01-14  quark-server.log.2022-12-21  quark-server.log.2023-02-01
driver_1670905747764.jstack quark-metastore.log.2022-12-27  quark-metastore.log.2023-02-07  quark-server.audit.2023-01-15  quark-server.log.2022-12-22  quark-server.log.2023-02-02
driver_1670905753618.jstack quark-metastore.log.2022-12-28  quark-metastore.log.2023-02-08  quark-server.audit.2023-01-16  quark-server.log.2022-12-23  quark-server.log.2023-02-03
driver_1670905759610.jstack quark-metastore.log.2022-12-29  quark-metastore.log.2023-02-09  quark-server.audit.2023-01-17  quark-server.log.2022-12-24  quark-server.log.2023-02-04
driver_1670907230244.jstack quark-metastore.log.2022-12-30  quark-metastore.log.2023-02-10  quark-server.audit.2023-01-18  quark-server.log.2022-12-25  quark-server.log.2023-02-05
driver_1670908241397.jstack quark-metastore.log.2022-12-31  quark-metastore.log.2023-01-01  quark-server.audit.2023-01-19  quark-server.log.2022-12-26  quark-server.log.2023-02-06
driver_1670908250221.jstack quark-metastore.log.2023-01-01  quark-server.audit.2022-12-09  quark-server.audit.2023-01-20  quark-server.log.2022-12-27  quark-server.log.2023-02-07
driver_167090812525.jstack quark-metastore.log.2023-01-02  quark-server.audit.2022-12-10  quark-server.audit.2023-01-21  quark-server.log.2022-12-28  quark-server.log.2023-02-08
driver_167090829349.jstack quark-metastore.log.2023-01-03  quark-server.audit.2022-12-11  quark-server.audit.2023-01-22  quark-server.log.2022-12-29  quark-server.log.2023-02-09
driver_1670912749885.jstack quark-metastore.log.2023-01-04  quark-server.audit.2022-12-12  quark-server.audit.2023-01-23  quark-server.log.2022-12-30  user_info_tab.txt
driver_1670915947173.jstack quark-metastore.log.2023-01-05  quark-server.audit.2022-12-13  quark-server.audit.2023-01-24  quark-server.log.2022-12-31
driver_1670919336147.jstack quark-metastore.log.2023-01-06  quark-server.audit.2022-12-14  quark-server.audit.2023-01-25  quark-server.log.2023-01-01
driver_1670919718154.jstack quark-metastore.log.2023-01-07  quark-server.audit.2022-12-15  quark-server.audit.2023-01-26  quark-server.log.2023-01-02
driver_1670924074518.jstack quark-metastore.log.2023-01-08  quark-server.audit.2022-12-16  quark-server.audit.2023-01-27  quark-server.log.2023-01-03
```

另外TDH服务保留日志的时间比较久、日志文件数往往是有多个，因此需要注意日志文件中的时间。

8.2.2. 示例2.误操作，社区开发版容器停止运行，如何恢复

【问题描述】

社区开发版使用时，因为误操作，使用docker attach <container id> 命令进入容器，并使用exit退出容器，导致容器退出并停止，集群停止运行。

【解决方法】

1. 在服务器端执行：docker ps -a，查看被关闭的容器ID。

```
[root@ce113 TDH-Client]# docker ps -a
CONTAINER ID        IMAGE               COMMAND             CREATED             STATUS              PORTS
f224e7afb3d1      tdh-standalone:2023.5  "/usr/bin/boot.sh"  2 weeks ago        Up About an hour
```

2. 执行: `docker start <container ID>` , 启动该容器
3. 启动后, Manager恢复。进入Manager界面会发现此时各服务均不可用, 图标显示为灰色
4. 在恢复集群时需要优先启动基础服务[LicenseServer/Aquila/(Guardian)]
5. 基础服务恢复健康后即可回到集群首页, 一键启动产品服务即可
6. 点击服务右上方“启动全部”按钮即可一键重启。



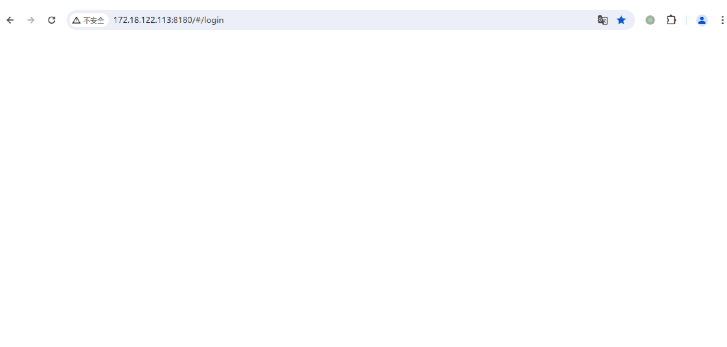
8.3. Manager及其他服务出现问题

8.3.1. Manager相关问题

8.3.1.1. 社区开发版

8.3.1.1.1. 8180 Manager页面打开后无反应

【问题描述】



【解决方法】

清楚缓存, 重新刷新

8.3.1.1.2. 8180 Manager页面显示拒绝访问

【问题描述】



【问题根因】

在启动新的容器时，环境还有正在运行的其他容器，端口未释放，导致启动新容器时端口冲突。即使后续有停止其他容器，但是此时manager自启动已失败，所以无法访问。

【解决方法】



在启动新容器之前，请再次务必确保环境上没有正在运行的其他开发版产品容器。

1. 进入容器

```
docker ps
docker exec -it <containerid> bash
```

2. 查看manager服务状态

```
systemctl status transwarp-manager-db
```

3. 重启manager服务

```
systemctl restart transwarp-manager-db
```

8.3.1.2. 社区版

8.3.1.2.1. TDH Manager登录不上

【用户问题】

用户名密码验证码都是正确的，但是点了登录只会刷新验证码。

【解决方法】

需要检查一下磁盘是不是满了。

8.3.1.2.2. 误操作，不小心删除了manager中的mysql，怎么办

【问题描述】

误操作不小心删除了

```

export SERVER_PRESTO_PLUGIN=mysql
您在 /var/spool/mail/root 中有新邮件
[root@hadoop01 ~]# rm -rf /etc/selinux/targeted/active/modules/100/mysql /var/lib/transwarp-manager/master/data/mysql/mysql /var/log/transwarp-manag
er/master/mysql
您在 /var/spool/mail/root 中有新邮件
[root@hadoop01 ~]# cd /var/log/transwarp-manager/master/
您在 /var/spool/mail/root 中有新邮件
    
```

【解决方法】

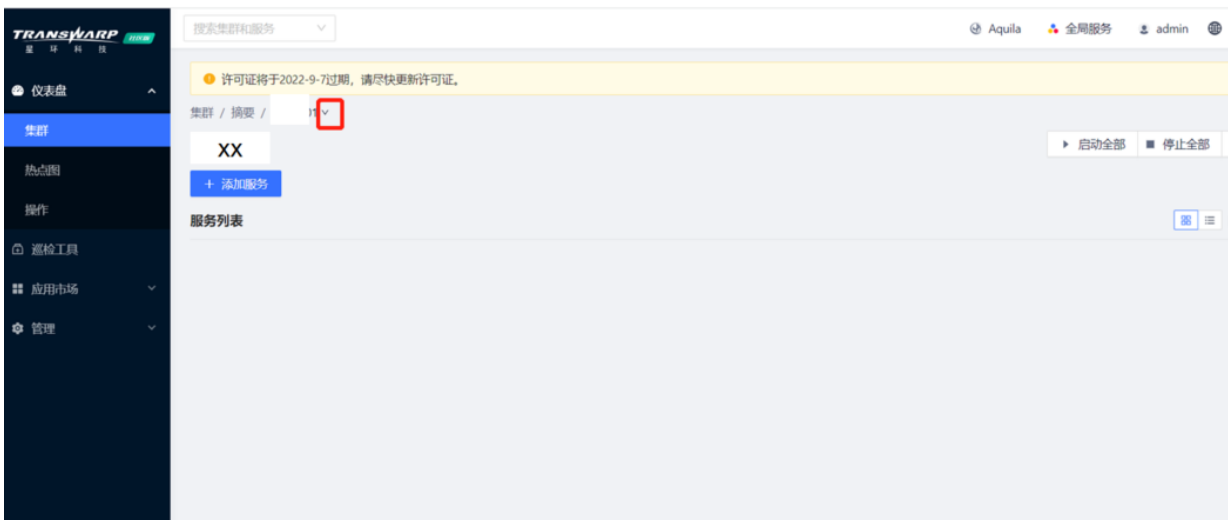
此处存储着manager的数据，如果不小心删掉了需要重新创建集群。

解决步骤：

把/etc/transwarp-manager/master/db.properties删掉，然后重启manager服务。

```
rm -rf /etc/transwarp-manager/master/db.properties
```

重启后可能会进入8179，访问8180即可。点击红框框进行添加集群。



8.3.2. Yarn相关问题

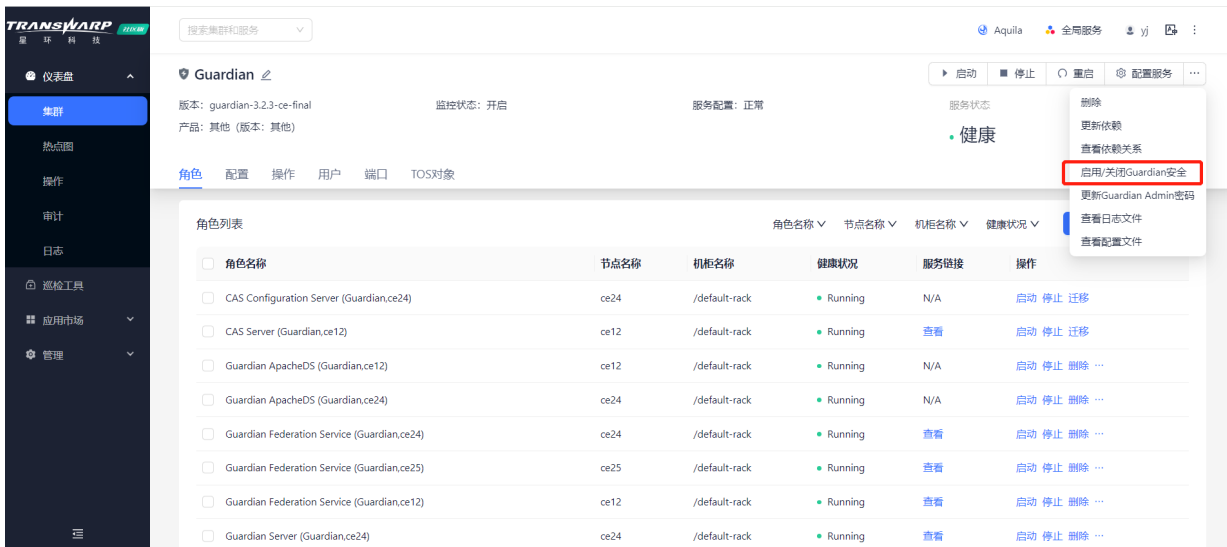
8.3.2.1. 启动yarn服务失败，显示“failure to login:for principle 以及 cannot locate KDC”

【核心原因】

这种时候通常是因为之前单独开关某个服务的安全顺序出现问题，没有注意依赖关系。

【解决方法】

重新在guardian页面统一开关安全，完成后重新尝试启动yarn服务。



8.3.2.2. 报错显示Invalid resource request...

【报错信息】

```

at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
at org.apache.hadoop.yarn.ipc.RPCUtil.instantiateYarnException(RPCUtil.java:53)
at org.apache.hadoop.yarn.ipc.RPCUtil.instantiateYarnException(RPCUtil.java:75)
at org.apache.hadoop.yarn.ipc.RPCUtil.unwrapAndThrowException(RPCUtil.java:116)
at org.apache.hadoop.yarn.api.impl.pb.client.ApplicationClientProtocolPBClientImpl.submitApplication(ApplicationClientProtocolPBClientImpl.java:297)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:422)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invokeMethod(RetryInvocationHandler.java:165)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invoke(RetryInvocationHandler.java:157)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invokeOnce(RetryInvocationHandler.java:95)
at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:359)
at com.sun.proxy.$Proxy15.submitApplication(Unknown Source)
at org.apache.hadoop.yarn.client.api.impl.YarnClientImpl.submitApplication(YarnClientImpl.java:384)
at org.apache.hadoop.mapred.ResourceMgrDelegate.submitApplication(ResourceMgrDelegate.java:299)
at org.apache.hadoop.mapred.YARNRunner.submitJob(YARNRunner.java:330)
... 17 more
Caused by: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.yarn.exceptions.InvalidResourceRequestException): Invalid resource request, requested resource type=<memory-mb> < 0 or greater than maximum allowed allocation. Requested resource=<memory:1536, vCores:1>, maximum allowed allocation=<memory:1024, vCores:2>. please note that maximum allowed allocation is calculated by scheduler based on maximum resource of registered NodeManagers, which might be less than configured maximum allocation=<memory:98304, vCores:96>
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.throwInvalidResourceException(SchedulerUtils.java:420)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.checkResourceRequestAgainstAvailableResource(SchedulerUtils.java:462)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.validateResourceRequest(SchedulerUtils.java:311)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:259)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:223)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.validateAndCreateResourceRequest(RMAppManager.java:559)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.createAndPopulateNewRMApp(RMAppManager.java:407)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.submitApplication(RMAppManager.java:342)
at org.apache.hadoop.yarn.server.resourcemanager.ClientRMService.submitApplication(ClientRMService.java:678)
at org.apache.hadoop.yarn.api.impl.pb.service.ApplicationClientProtocolPBServiceImpl.submitApplication(ApplicationClientProtocolPBServiceImpl.java:277)
at org.apache.hadoop.yarn.proto.ApplicationClientProtocol$ApplicationClientProtocolService$2.callBlockingMethod(ApplicationClientProtocol.java:563)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtobufRpcInvoker.call(ProtobufRpcEngine.java:527)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1036)
at org.apache.hadoop.ipc.Server$RpcCall.run(Server.java:1055)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:2145)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2953)

at org.apache.hadoop.ipc.Client.getRpcResponse(Client.java:1582)
at org.apache.hadoop.ipc.Client.call(Client.java:1528)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:233)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:118)
at com.sun.proxy.$Proxy14.submitApplication(Unknown Source)
at org.apache.hadoop.yarn.api.impl.pb.client.ApplicationClientProtocolPBClientImpl.submitApplication(ApplicationClientProtocolPBClientImpl.java:294)
... 30 more
Data generation failed.
    
```

【解决方法】

```
Requested resource=<memory:1536, vCores:1>, maximum allowed allocation=<memory:1024, vCores:2>,
```

yarn的node manager内存资源给少了，可以到manager yarn服务的配置列表调整yarn.nodemanager.resource.memory-mb这个参数

8.3.3. Quark相关问题

8.3.3.1. Quark服务状态异常

【问题描述】

quark服务安装完成，状态健康，但是一分钟后出现重启，大概重启了两次，然后自动关闭了

```
Failed to create scratch, org.apache.hadoop.security.AccessControlException: Permission denied: user=hive, access=WRITE, inode="/":hdfs:hbase:drwxr-xr-x
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:255)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1874)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1858)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkAncestorAccess(FSDirectory.java:1817)
at org.apache.hadoop.hdfs.server.namenode.FSDirMkdirOp.mkdirs(FSDirMkdirOp.java:59)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.mkdirs(FSNamesystem.java:3340)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.mkdirs(NameNodeRpcServer.java:1189)
at org.apache.hadoop.hdfs.protocolPB.ClientNameNodeProtocolServerSideTranslatorPB.mkdirs(ClientNameNodeProtocolServerSideTranslatorPB.java:727)
at org.apache.hadoop.hdfs.protocol.proto.ClientNameNodeProtocolProtos$ClientNameNodeProtocol$2.callBlockingMethod(ClientNameNodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:527)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1036)
at org.apache.hadoop.ipc.Server$RpcCall.run(Server.java:1015)
at org.apache.hadoop.ipc.Server$RpcCall.run(Server.java:943)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:2145)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2953)

+ shift 1
+ '[' 0 -eq 0 ']'
+ set +x
no inceptor process alive
exit now
[root@ce12 ~]#
```

状态	描述	开始时间	结束时间	操作
SUCCESSFUL	生成依赖的服务的配置	2022-07-29 20:30:23	2022-07-29 20:30:23	查看日志
FAILED	生成 Collector 配置	2022-07-29 20:30:37	2022-07-29 20:30:37	重试 查看日志
SUCCESSFUL	创建数据库	2022-07-29 20:30:23	2022-07-29 20:30:23	查看日志
FAILED	配置Quark Metastore(ce12)	2022-07-29 20:30:23	2022-07-29 20:30:23	重试 查看日志
FAILED	配置Quark Metastore(ce24)	2022-07-29 20:30:23	2022-07-29 20:30:23	重试 查看日志
FAILED	配置Quark Server(ce12)	2022-07-29 20:30:37	2022-07-29 20:30:38	重试 查看日志
WAITING	配置Nucleon Executor(ce12)			查看日志
WAITING	配置Nucleon Executor(ce24)			查看日志
WAITING	配置Nucleon Executor(ce25)			查看日志
WAITING	通知 Aquila 重载配置文件			查看日志
初始化Quark1 0%				
状态	描述	开始时间	结束时间	操作
WAITING	在hdfs上创建目录			查看日志

【核心原因】

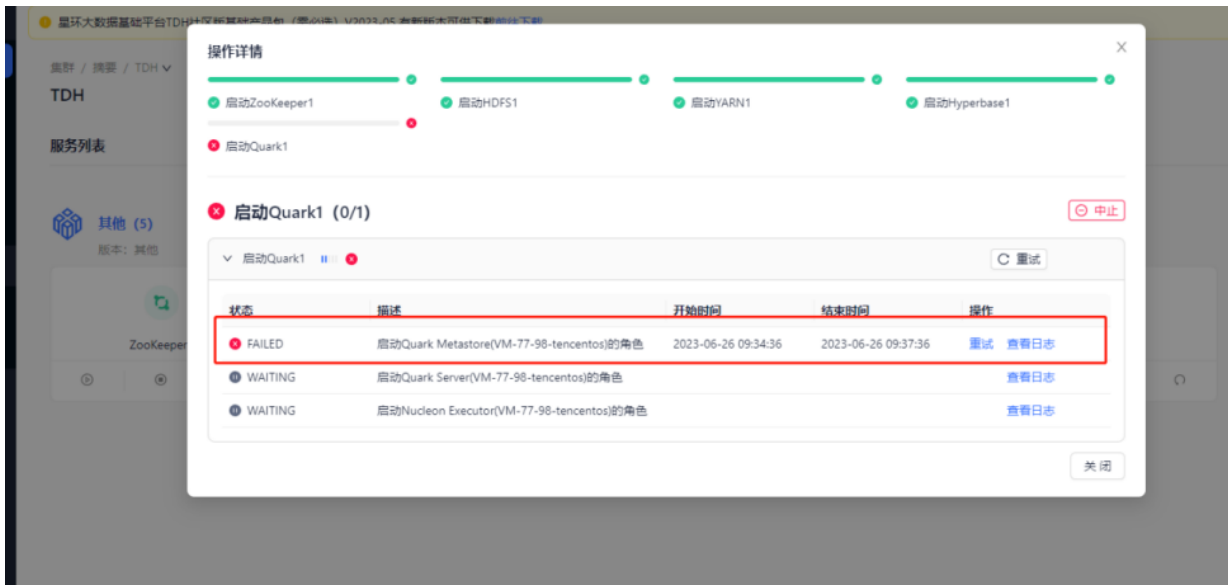
hdfs创建目录没执行，所以在安装的时候就失败了

【解决方法】

就是初始化quark1的时候，有错误，后面的步骤没有继续，重试即可

8.3.3.2. 启动Quark服务失败，无法启动集群角色

【报错情况】



【核心原因】

节点主机名配置错误



节点名称设置时不符合要求，须按照手册要求设定。

集群的节点名称不能重复，且必须符合DNS-1123规范，由数字、小写字母或“-”组成，不能包含大写字母，长度小于63。

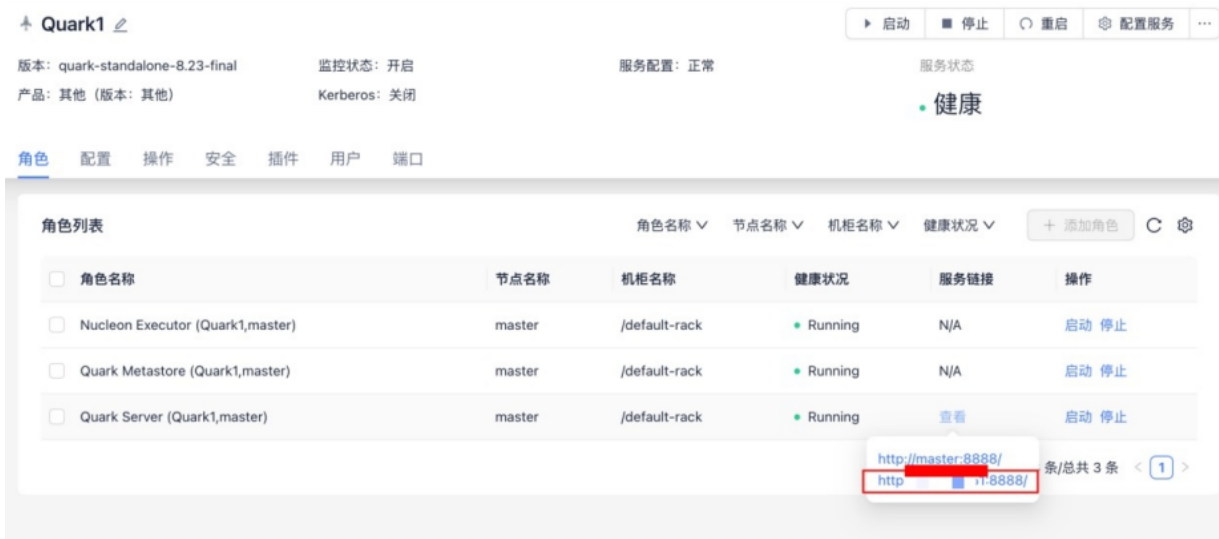
【解决方法】

修改后重试，可选择删除该容器重新启动一个容器的形式，重新安装产品服务。

8.3.3.3. 通过8888端口访问Quark Server，出现提醒：使用4040端口DBA Server服务

【用户问题】

咨询一下目前我们这边暂时在测试环境部署了单机社区开发版的TDH各个服务运行正常，但是我在尝试访问Quark Server服务跳转UI页面时出现提示需要使用标准DBA服务，请问这种情况是否正常？

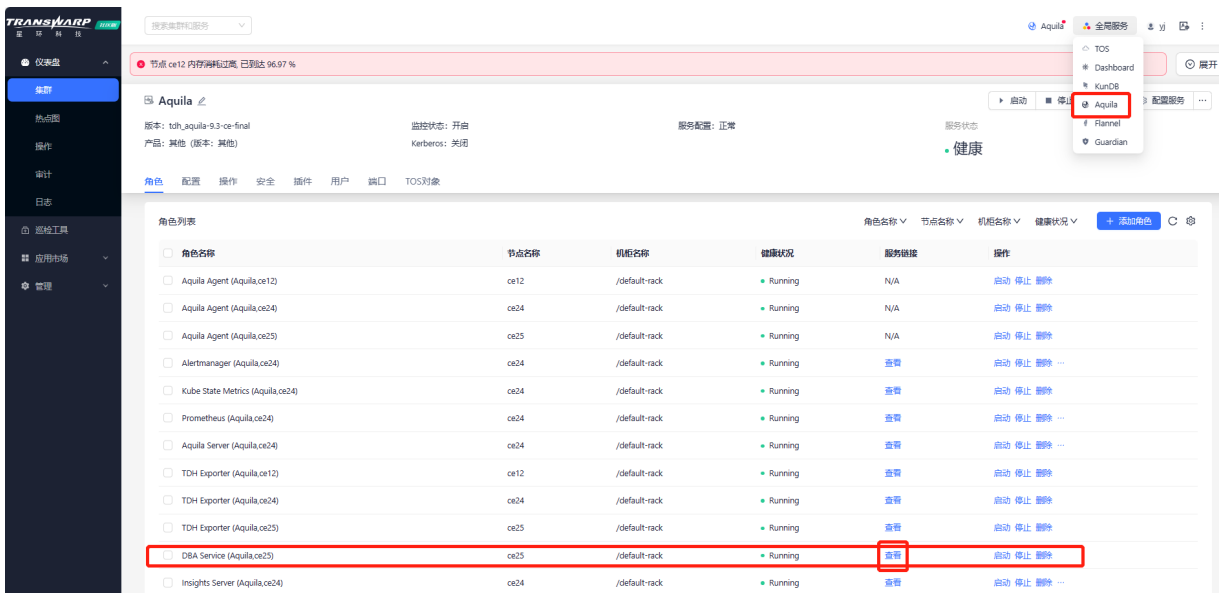
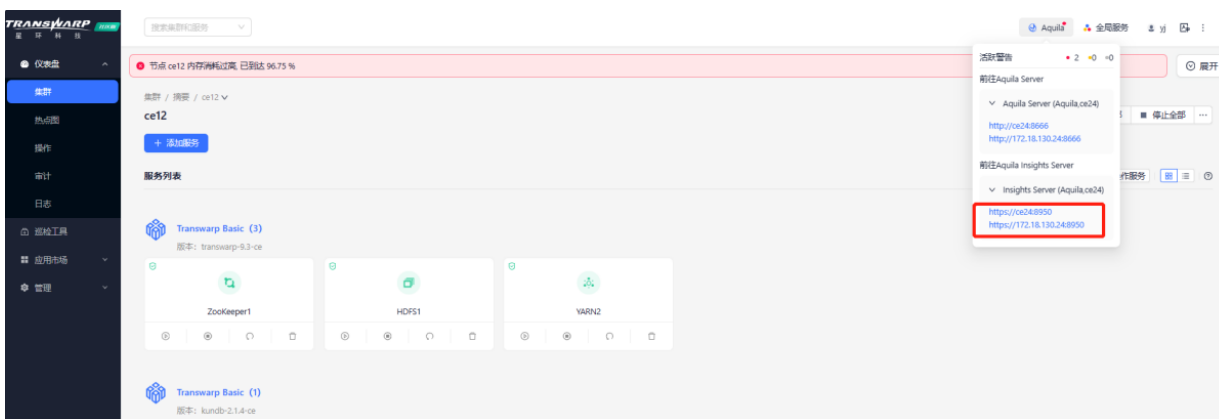


← → ↻ ⚠ 不安全 1:8888/jobs

Please use standard DBA Service or consult DBA Service Staff for how to enable old 4040

【问题答复】

您好，这是正常的，您可以通过点击右上角的insight server (DBA)服务查看，也可进入全局服务的Aquila服务详情页面查看DBA Service角色的相关链接进行访问：



8.3.4. Guardian相关问题

8.3.4.1. Guardian开启安全，开启ZK服务安全时报错，报错显示：error.nodeHostNotFound

【报错情况】

生成kerberos配置文件

开始时间: 2023-06-12 14:56:31

结束时间: 2023-06-12 14:56:32

操作详情:

↓ 导出日志

× 全屏

```

10 2023-06-12T06:56:32.048529 [Master] Copy Done
11 2023-06-12T06:56:32.048548 [Master] Distribute keytab to yun
12 2023-06-12T06:56:32.048553 [Master] Generate global configurations to manager nodes
13 2023-06-12T06:56:32.049933 [Master] The Task 287 run failed: java.lang.RuntimeException:
io.transwarp.manager.master.common.AppException: error.nodeHostNotFound
14     at io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper.runTask(TaskDriver.java:277)
15     at
io.transwarp.manager.master.operation.execution.TaskDriver$TaskRunHelper$$FastClassBySpringCGLIB$$e353057f.invoke(<generated>)
16     at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
17     at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:793)
18     at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
19     at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:763)
20     at org.springframework.aop.interceptor.AsyncExecutionInterceptor.lambda$invoke$0(AsyncExecutionInterceptor.java:115)
21     at java.base/java.util.concurrent.FutureTask.run(FutureTask.java:264)
22     at java.base/java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1128)
23     at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628)
24     at java.base/java.lang.Thread.run(Thread.java:829)

```

取消

重试

【核心原因】

host文件中的IP与实际IP不一致

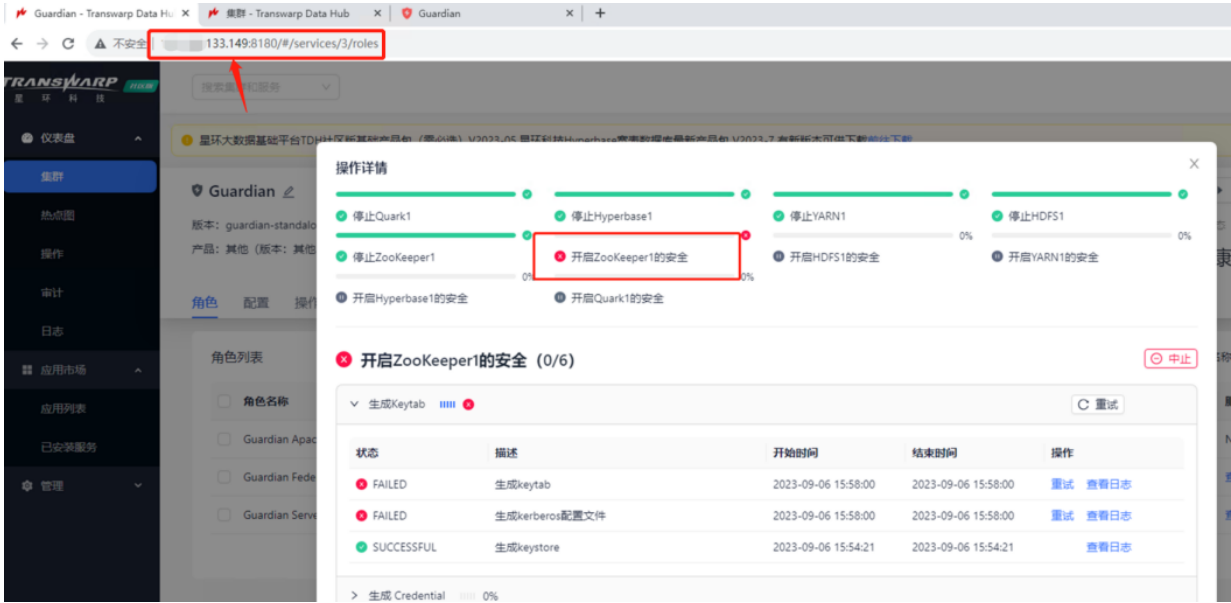
【解决方法】

检查hostname及/etc/hosts文件，不一致修改后重试即可

```

[root@slave0 ~]#
[root@slave0 ~]#
[root@slave0 ~]# hostname
slave0
[root@slave0 ~]#
[root@slave0 ~]# cat /etc/hosts
127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost localhost.localdomain localhost6 localhost6.localdomain6
.133.140 master.baicdt.com master
.133.141 slave0
.133.142 slave1.baicdt.com slave1
bigdata03
bigdata31
bigdata35
bigdata24
bigdata32
bigdata35

```

8.3.4.2. 重启Guardian时通过TOS停止server失败，显示 “Guardian_server has no ready pod”

【问题描述】

重启Guardian时失败，步骤为“通过TOS停止server”，显示内容为” Guardian_server has no ready pod”



查看Guardian Server日志详情，显示如下：

Guardian_server 的pod在ce24节点处于Terminating状态：

```

2023-09-06T15:00:24.443824 [Master] Waiting 2 GUARDIAN_SERVER role(s) in Guardian to stop within 600 s
2023-09-06T15:10:24.511158 [Master] Latest health check result of roles:
VITAL_SIGN_CHECK DOWN at 2023-09-06T15:10:19.576

GUARDIAN_SERVER has no ready pod on ce12DAEMON_CHECK DOWN at 2023-09-06T15:10:18.924
GUARDIAN_SERVER has no Pod on ce12VITAL_SIGN_CHECK HEALTHY at 2023-09-06T15:10:19.576

GUARDIAN_SERVER on ce24 has pod guardian-server-guardian-5c49765679-rxcmx with readiness 1/1DAEMON_CHECK WARNING at 2023-09-06T15:10:18.924
GUARDIAN_SERVER on ce24 has pod guardian-server-guardian-5c49765679-rxcmx with status Terminating
2023-09-06T15:10:24.511340 [Master] Fail to run task remote part: java.lang.IllegalStateException: 2 GUARDIAN_SERVER role(s) in Guardian didn't stop within 600 s
    
```

通过查看TOS服务详情查看状态：

分类	检查时间	检查结果
DAEMON_CHECK	2023-09-06 17:29:54	WARNING

TOS_MASTER is running on ce25
 etcd --data-d is running as be14e3ce35923d9d9768b840663fd1653793e4fe64c664febaee73eae797e971
 kube-apiserver is running as 40df1fae67e67aeca79b65be68d77344c8c49b4407cfb4d9e4bb9aac10bfc3
 kube-controller is running as 268e4acbf3984ac168710ac14d07ebbad691644a141f7e9669d3ca486adfddee
 kube-scheduler is running as ac52b338f7d3f10a0a7005b24abd53f35e7a76227e725cf20b8288aed124577

TOS_MASTER is not running on ce24
 etcd --data-d has no running container
 kube-apiserver has no running container
 kube-controller has no running container
 kube-scheduler has no running container

TOS_REGISTRY are all healthy
 TOS_REGISTRY is running on ce12

VITAL_SIGN_CHECK	2023-09-06 17:29:51	WARNING
------------------	---------------------	---------

VITAL_SIGN_CHECK 2023-09-06 17:30:09 WARNING

```
Not all TOS_SLAVE are healthy
Connection check result of TOS Slave ce12 was HEALTHY
Connection check result of TOS Slave ce24 was DOWN
Connection check result of TOS Slave ce25 was HEALTHY

TOS_MASTER are all healthy
Connection check result of TOS Master ce12 was Healthy
Connection check result of TOS Master ce25 was Healthy
Connection check result of TOS Master ce24 was Healthy

TOS_REGISTRY are all healthy
Connection check result of Tos Registry ce12 was Healthy
```

常见pod的几种状态

- Pending: 这个状态意味着，Pod 的 YAML 文件已经提交给了 Kubernetes，API 对象已经被创建并保存在 Etcd 当中。但是这个 Pod 还没有被调度成功，最常见的原因比如 Pod 中某个容器启动不成功
- Running: 这个状态下，Pod 已经调度成功。也就是它包含的容器都已经创建成功，并且至少有一个正在运行中
- Succeeded: 这个状态意味着，Pod 里的所有容器都正常运行成功并退出了。这种情况在运行一次性任务时最为常见
- Failed: 这个状态下，Pod 里至少有一个容器以不正常的状态退出。这个状态出现时，你得想办法 Debug 这个容器，比如查看 Pod 的事件和日志。
- Unknown: 这是一个异常状态，意味着 Pod 的状态不能集群检测到，这很有可能是主从节点（Master 和 Kubelet）间的通信出现了问题。
- Terminating: 在节点处于“NotReady”状态时，deployment控制器会迁移节点上的容器实例，并将节点上运行的pod置为“Terminating”状态。待节点恢复后，处于“Terminating”状态的pod会自动删除。偶现部分pod（实例）一直处于“Terminating”状态，发现这部分的pod没有得到重新调度，不能提供服务。

然后，在服务器端，查看集群中pod的状态，发现guardian的一个pod处于Terminating，同时排除其他pod报错。到目前为止，需要对其进行强制删除，并重启该pod以恢复Guardian运行。

```
[root@ce12 guardian]# kubectl get pods -owide
NAME                                READY   STATUS             RESTARTS   AGE   IP              NODE     NOMINATED NODE   READINESS GATES
alertmanager-aquila-5c745b4cfb-j87fc 1/1     Running            0          14d   172.18.130.12  ce12    <none>            <none>
aquila-agent-aquila-57d6cbbc96-kngwc 1/1     Running            0          14d   172.18.130.25  ce25    <none>            <none>
aquila-agent-aquila-57d6cbbc96-rsb26 1/1     Running            0          14d   172.18.130.12  ce12    <none>            <none>
aquila-agent-aquila-57d6cbbc96-x76kq 1/1     Running            0          14d   172.18.130.24  ce24    <none>            <none>
aquila-server-aquila-6fff586779-pghrj 1/1     Running            0          14d   172.18.130.12  ce12    <none>            <none>
dba-service-aquila-5f44b6846f-285nh 1/1     Running            0          14d   172.18.130.12  ce12    <none>            <none>
guardian-federation-guardian-586bdfdd4f-nxk2b 1/1     Running            1          36d   172.18.130.12  ce12    <none>            <none>
guardian-federation-guardian-586bdfdd4f-x6prx 1/1     Running            1          36d   172.18.130.24  ce24    <none>            <none>
guardian-server-guardian-5c49765679-rxcmx 1/1     Terminating      1          36d   172.18.130.24  ce24    <none>            <none>
hadoop-hdfs-datanode-hdfs1-74958/d67c-765vk 1/1     Running            5          63d   172.18.130.12  ce12    <none>            <none>
hadoop-hdfs-datanode-hdfs1-74958/d67c-tg4tq 1/1     Running            2          63d   172.18.130.24  ce24    <none>            <none>
hadoop-hdfs-datanode-hdfs1-74958/d67c-tkz2w 1/1     Running            4          63d   172.18.130.25  ce25    <none>            <none>
hadoop-hdfs-https-hdfs1-5854446dc4-5hzk4 1/1     Running            5          63d   172.18.130.12  ce12    <none>            <none>
hadoop-hdfs-journalnode-hdfs1-584d9c5586-b4mxz 1/1     Running            2          63d   172.18.130.24  ce24    <none>            <none>
```

定位到pod以后，使用kubectl logs <pod-names>查看pod的日志信息。此时看到无法查看日志。

```
[root@ce12 ~]# kubectl logs guardian-server-guardian-5c49765679-rxcmx
Error from server (NotFound): the server could not find the requested resource ( pods/log guardian-server-guardian-5c49765679-rxcmx)
```

随后，通过systemctl status命令查看是否存在失败的服务，发现有1个失败的服务。

```

[unknown] operation succeeded.
[root@ce12 ~]# systemctl status
● ce12
  State: degraded
    Jobs: 0 queued
  Failed: 1 units
  Since: Wed 2023-08-16 15:57:13 CST; 3 weeks 0 days ago
  CGroup: /
├─1 /usr/lib/systemd/systemd --switched-root --system --deserialize 22
├─docker
│   └─87e6dba7395b9cf381cee0a179cbbbed7c7efa6fc9bd7630ae75136ad08ece82b
│       └─272414 bash
├─kubepods
└─podf14e8cf1-71ed-435c-beb3-02e51a8f479f

```

通过systemctl --failed命令查看失败的服务

```

[root@ce12 ~]# systemctl --failed
UNIT                                LOAD    ACTIVE SUB    DESCRIPTION
● docker.socket loaded failed failed Docker Socket for the API

```

判断是由于docker.socket服务出现问题，导致重启服务时，无法为服务分配新的pod。

【修复方法】

```

# systemctl stop docker
# kill $(ps -aux|grep docker |awk '{print $2}')
# kill $(ps -ef | grep docker-containerd-shim | awk '{ print $2 }')
# rm -rf /var/lib/docker/*
# docker-storage-setup --reset (可跳过)
# docker-storage-setup (可跳过)
# systemctl restart docker
# docker load -i /etc/tos/conf/tos.tar.gz

```

```

[root@ce24 ~]# kill $(ps -aux|grep docker |awk '{print $2}')
-bash: kill: (872412) - No such process
[root@ce24 ~]# kill $(ps -ef | grep docker-containerd-shim | awk '{ print $2 }')
-bash: kill: (872553) - No such process
[root@ce24 ~]# rm -rf /var/lib/docker/*

```

```

[root@ce24 ~]# systemctl restart docker
[root@ce24 ~]# docker load -i /etc/tos/conf/tos.tar.gz
cdb3f9544e4c: Loading layer [=====>] 58.44MB/58.44MB
48eb6d556e52: Loading layer [=====>] 7.871MB/7.871MB
Loaded image: transwarp/nvidia-device-plugin:tos-2.1.2
8d3ac3489996: Loading layer [=====>] 5.866MB/5.866MB
d1b86b32010c: Loading layer [=====>] 39.12MB/39.12MB
b55875d1c702: Loading layer [=====>] 39.12MB/39.12MB
Loaded image: transwarp/static-node-scheduler-extender:v0.1.0

```

```

[root@ce24 ~]# docker-storage-setup --reset
-bash: docker-storage-setup: command not found
[root@ce24 ~]# docker-storage-setup
-bash: docker-storage-setup: command not found
[root@ce24 ~]# systemctl restart docker
[root@ce24 ~]# systemctl status docker
● docker.service - Docker Application Container Engine
  Loaded: loaded (/usr/lib/systemd/system/docker.service; enabled; vendor preset: disabled)
  Active: active (running) since Wed 2023-09-06 21:07:00 CST; 9s ago
  Docs: http://docs.docker.com
  Process: 886104 ExecStartPre=/bin/sh -ec ip link delete docker0 || true (code=exited, status=0/SUCCESS)
  Main PID: 886127 (dockerd)
  Tasks: 21
  Memory: 217.2M
  CGroup: /machine.slice/docker.service
├─886127 /usr/bin/dockerd --log-opt max-size=500m --log-opt max-file=2 --live-restore --iptables=false
└─886159 docker-containerd --config /var/run/docker/containerd/containerd.toml

```

恢复docker服务以后，在Manager界面进行TOS服务的重启，重启成功。随后分别进行Aquila和Guardian服务的重启，重启成功。最后一键重启其余服务，集群恢复运行。

【修复方法说明】

docker创建的容器是无状态的，一些状态信息（例如数据、配置文件等）其实都已经挂载进宿主机或者持

久化到其他地方；这个时候可以通过上面的修复方法，关闭docker相关容器和进程，删除/var/lib/docker/下面容器相关的无状态信息；然后重启docker和kubelet，这个时候k8s会自动将之前的容器调度起来；

8.3.4.3. 如果忘记或需要重置Guardian中管理员密码，如何操作？

【环境描述】

TDH社区版-V2023-5版本

TDH社区开发版-V2023-5版本

Guardian-V3.2.3、3.2.5版本

后面将针对用户忘记Guardian admin (KAdmin) 密码或需要重置密码的情况。如果出现无法使用“开启/关闭Guardian安全”和无法访问Guardian Server的情况，请参考后续解决步骤。

【问题描述】

无法使用“开启/关闭Guardian安全”，显示如下



此时需要确认密码是否无误，无误后检查是否可以访问Guardian Server



此时发现Server页面打不开。

Server页面打不开的情况下需要先查看是否已修改本地hosts文件将主机名与Manager相关服务页面的IP地址进行关联。例如 Windows 平台下，该文件的路径为

C:\Windows\System32\drivers\etc\hosts。

此时发现未做关联，添加后即可访问，但是输入相关密码后依然无法登录。此时则判定可能是密码出了问题。



那么如何重置密码?

【解决步骤】

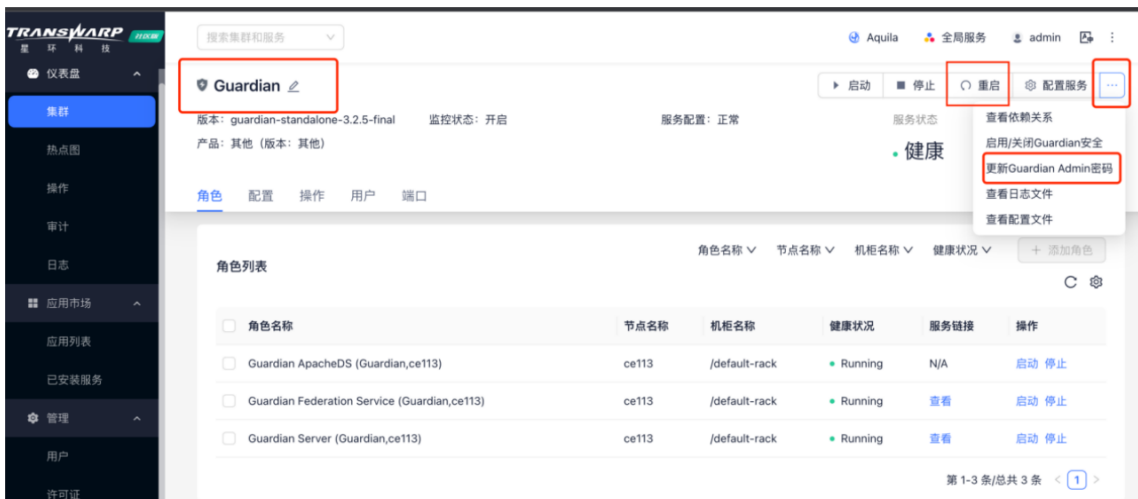
1. 版本: 社区版-V23-5版本

a. 执行 `kubectl grep po -owide | grep guardian` 命令进入 guardian pod, 执行重置密码的脚本文件

```
[root@sec11 ~]# kubectl get po -owide | grep guardian
cas-server-guardian-6d996f7cct-vn8wc      1/1      Running 2      13d      172.18.130.11      sec11      <none>      <none>
guardian-apacheds-guardian-7d55b76db5-9mh9j  1/1      Running 2      13d      172.18.130.14      sec14      <none>      <none>
guardian-apacheds-guardian-7d55b76db5-nzdkj  1/1      Running 1      13d      172.18.130.11      sec11      <none>      <none>
guardian-federation-guardian-f68775c9b-dfb7b  1/1      Running 1      13d      172.18.130.14      sec14      <none>      <none>
guardian-federation-guardian-f68775c9b-ghcsg  1/1      Running 1      13d      172.18.130.11      sec11      <none>      <none>
guardian-server-guardian-766c5bcd6d-8dq17    1/1      Running 1      13d      172.18.130.14      sec14      <none>      <none>
guardian-server-guardian-766c5bcd6d-w4t5c    1/1      Running 1      13d      172.18.130.11      sec11      <none>      <none>
[root@sec11 ~]# kubectl exec -it guardian-server-guardian-766c5bcd6d-w4t5c bash
[root@sec11 ~]# cd /usr/lib/guardian/
[root@sec11 guardian]# ls
default-conf  guardian-server-boot-guardian-3.2.6.jar  lib  scripts
[root@sec11 guardian]# cd scripts/
[root@sec11 scripts]# ls
bootstrap.sh          init-krb5-conf.sh
ehcache.xml          init-ldap-connection.sh
import-apacheds-data.sh  init-log4j-properties.sh
init-application-properties.sh  kadmin.guardian
init-ehcache-conf.sh   resetpwd
init-guardian-site.sh  wait-for-it.sh
[root@sec11 scripts]# resetpwd --help
[root@sec11 scripts]# resetpwd --help
usage: ResetPwdShell
       -newPassword <arg>      New Password
       -username <arg>         The User Name
[root@sec11 scripts]#
```

执行 `sh resetpwd -newPassword <新密码> -username admin`

b. 执行完成后, Guardian admin的密码已经修改, 接下来需要去Manager界面更新密码。





点击“更新Guardian Admin密码”后，在出现的弹窗中输入刚刚重设的密码，点击确定。

然后界面重启Guardian服务即可。

2. 版本：社区开发版-V23-5版本

- a. 在服务器端执行：docker ps ，查看正在运行的容器ID。

```
[root@ce113 ~]# docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED        STATUS        PORTS        NAMES
c224e7afb3d1  tdh-standalone:2023.5  "/usr/bin/boot.sh"     2 weeks ago   Up 2 hours   -           laughing_noether
```

- b. 执行：docker exec -it <container ID> bash ，进入容器内部。

```
[root@ce113 ~]# docker exec -it c224e7afb3d1 bash
[root@ce113 /]#
```

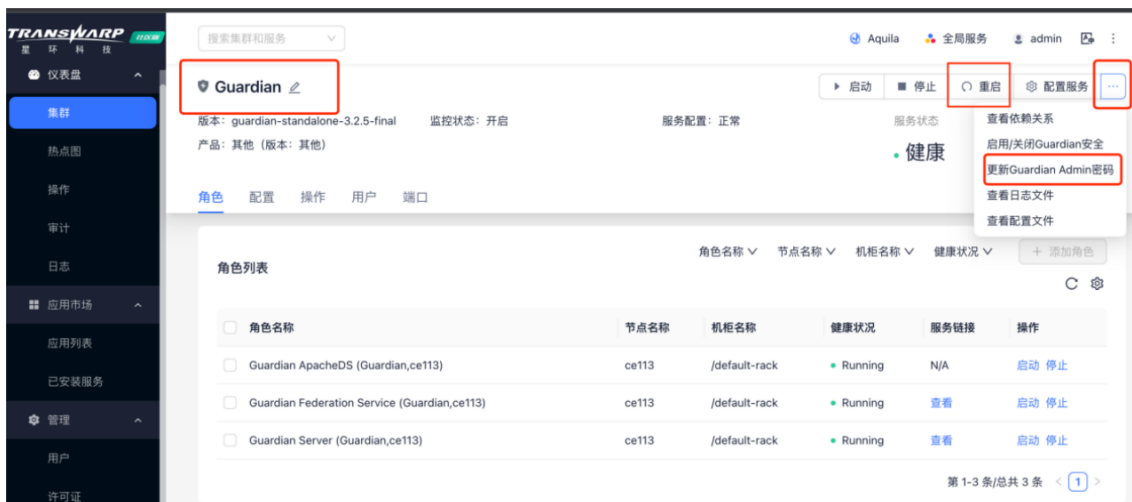
- c. 进入/lib/guardian/scripts目录下，可以看到该目录下有名为“resetpwd”的脚本文件，其执行参数如下图所示。接下来就可以通过执行该脚本重置Guardian的admin用户密码。

```
[root@ce113 scripts]# pwd
/lib/guardian/scripts
[root@ce113 scripts]# sh resetpwd --help
usage: ResetPwdShell
       -newPassword <arg>   New PassWord
       -username <arg>     The User Name
```

- d. 通过执行脚本文件实现密码修改，在此修改用户为admin，您可设定新的密码，命令如下所示；执行：sh resetpwd -newPassword <新密码> -username admin

```
[root@ce113 scripts]# sh resetpwd -newPassword . -username admin
2023-06-30 09:21:07,288 INFO pool.DruidDataSource: {dataSource-1} inited
2023-06-30 09:21:08,210 INFO utils.ResetPwdShell: Password of user admin has been successfully changed
```

- e. 执行完成后，Guardian admin的密码已经修改，接下来需要去Manager界面更新密码。



点击“更新Guardian Admin密码”后，在出现的弹窗中输入刚刚重设的密码，点击确定。然后重启Guardian服务。

此时，已全部完成Guardian Admin密码的重设步骤，可以去Guardian Server界面尝试登录，验证新密码是否生效。

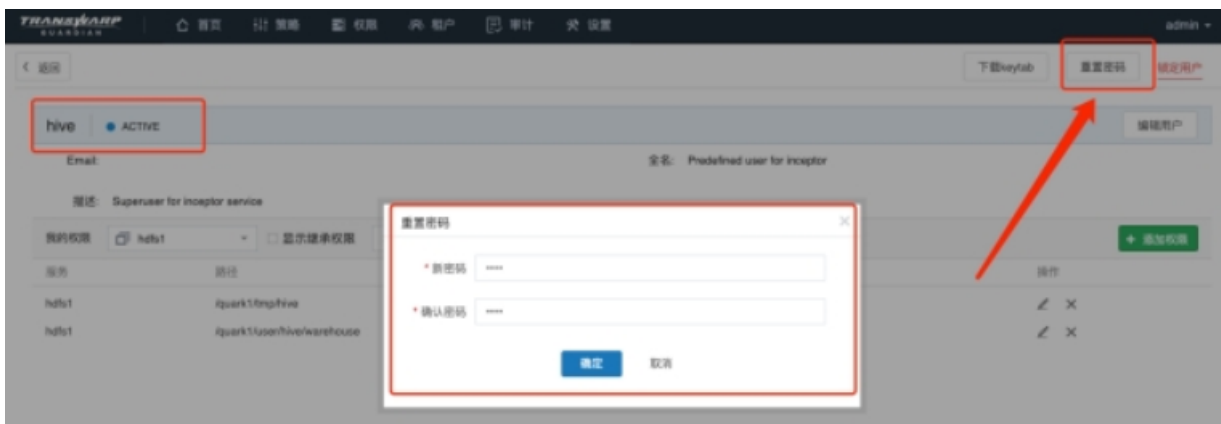
8.3.4.4. 如何在Guardian上重置租户密码?

【用户问题】

请问下密码可在Guardian界面重置密码获取。这一步怎么操作的，找不到重置的地方

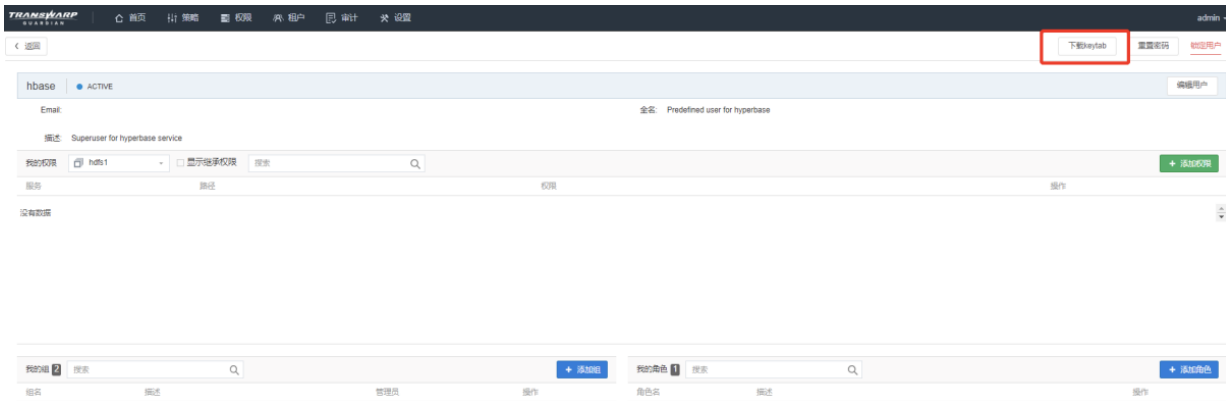
【问题答复】

打开Guardian Server页面后，点击租户—选择您要使用的用户 - 重置密码即可。



8.3.4.5. 社区开发版如何在guardian上下载hbase租户的keytab

【解决方法】



可以在上图所示位置进行下载，或者也可以在后台挂载目录下进行查看，比如开发版容器挂载的目录路径是/home/tdh，那么就在/home/tdh/conf/hyperbase1/conf。目录下有hyperbase.keytab。

8.3.5. HFDS相关问题

8.3.5.1. 初始化 HDFS 预定义监控信息 启动失败

【报错信息】

状态	描述	开始时间	结束时间	操作
SUCCESSFUL	通过TOS启动hdfs-https	2023-05-25 15:28:47	2023-05-25 15:28:55	查看日志
SUCCESSFUL	为Data Node(yq-ll-centos7-23-31)添加TOS标签	2023-05-25 15:28:55	2023-05-25 15:28:55	查看日志
SUCCESSFUL	为Data Node(yq-ll-centos7-23-32)添加TOS标签	2023-05-25 15:28:55	2023-05-25 15:28:55	查看日志
SUCCESSFUL	为Data Node(yq-ll-centos7-23-33)添加TOS标签	2023-05-25 15:28:55	2023-05-25 15:28:55	查看日志
SUCCESSFUL	为Data Node(yq-ll-centos7-23-37)添加TOS标签	2023-05-25 15:28:55	2023-05-25 15:28:55	查看日志
SUCCESSFUL	通过TOS启动hdfs-datanode	2023-05-25 15:28:55	2023-05-25 15:29:19	查看日志
FAILED	初始化 HDFS 预定义监控信息	2023-05-25 15:29:19	2023-05-25 15:29:20	重试 查看日志

```

2023-05-25T15:29:19.977 [Master] ===== Task 647 start to run. =====
2023-05-25T15:29:19.978 [Master] Starting task local part ...
2023-05-25T15:29:20.006 [Master] yq-ll-centos7-23-31 send post request to http://192.168.23.31:8666/api/metrics/dashboards/serviceDashboard/init to Initialize predefined dashboards
2023-05-25T15:29:20.024 [Master] The Task 647 run failed: java.lang.RuntimeException: org.apache.http.conn.HttpHostConnectException: Connect to 192.168.23.31:8666 [/192.168.23.31] failed: Connection refused (Connection refused)
    at io.transwarp.manager.master.operation.execution.TaskDriverTaskRunner.runTask(TaskDriver.java:224)
    at io.transwarp.manager.operation.execution.TaskDriverTaskRunner$FastClassBySpringAOPWeld.invoke()
    at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
    at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:783)
    at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
    at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:753)
    at org.springframework.aop.interceptor.AsyncExecutionInterceptor.lambda$invoke$0(AsyncExecutionInterceptor.java:115)
    at java.util.concurrent.FutureTask.run(FutureTask.java:266)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
    at java.lang.Thread.run(Thread.java:748)
Caused by: org.apache.http.conn.HttpHostConnectException: Connect to 192.168.23.31:8666 [/192.168.23.31] failed: Connection refused (Connection refused)
    at org.apache.http.impl.conn.DefaultHttpClientConnectionOperator.connect(DefaultHttpClientConnectionOperator.java:156)
    at org.apache.http.impl.conn.PoolingHttpClientConnectionManager.connect(PoolingHttpClientConnectionManager.java:376)
    at org.apache.http.impl.execchain.MainClientExec.establishRoute(MainClientExec.java:393)
    at org.apache.http.impl.execchain.MainClientExec.execute(MainClientExec.java:234)
    at org.apache.http.impl.execchain.ProtocolExec.execute(ProtocolExec.java:186)
    at org.apache.http.impl.execchain.RetryExec.execute(RetryExec.java:89)
    at org.apache.http.impl.execchain.RedirectExec.execute(RedirectExec.java:110)
    at org.apache.http.impl.client.InternalHttpClient.doExecute(InternalHttpClient.java:185)
    at org.apache.http.impl.client.CloseableHttpClient.execute(CloseableHttpClient.java:72)
    at org.apache.http.impl.client.CloseableHttpClient.execute(CloseableHttpClient.java:221)
    at org.apache.http.impl.client.CloseableHttpClient.execute(CloseableHttpClient.java:165)
    at org.apache.http.impl.client.CloseableHttpClient.execute(CloseableHttpClient.java:140)
    at io.transwarp.manager.common.util.HttpUtils.post(HttpUtils.java:50)
    at io.transwarp.manager.operation.execution.localrunner.AbstractTaskLocalRunner.executeHttpPost(AbstractTaskLocalRunner.java:668)
    at io.transwarp.manager.operation.execution.localrunner.InitPredefinedOpMonitorInfoTaskLocalRunner.initDashboard(InitPredefinedOpMonitorInfoTaskLocalRunner.java:199)
    at io.transwarp.manager.operation.execution.localrunner.InitPredefinedOpMonitorInfoTaskLocalRunner.execute(InitPredefinedOpMonitorInfoTaskLocalRunner.java:87)
    at io.transwarp.manager.operation.execution.localrunner.InitPredefinedOpMonitorInfoTaskLocalRunner.execute(InitPredefinedOpMonitorInfoTaskLocalRunner.java:99)
    at io.transwarp.manager.operation.execution.localrunner.AbstractTaskLocalRunner.run(AbstractTaskLocalRunner.java:98)
    at io.transwarp.manager.operation.execution.TaskDriverTaskRunner.runTask(TaskDriver.java:224)
    ... 10 more
Caused by: java.net.ConnectException: Connection refused (Connection refused)
    at java.net.PlainSocketImpl.socketConnect(Native Method)
    at java.net.AbstractPlainSocketImpl.doConnect(AbstractPlainSocketImpl.java:350)
    at java.net.AbstractPlainSocketImpl.connectToAddress(AbstractPlainSocketImpl.java:206)
    at java.net.AbstractPlainSocketImpl.connect(AbstractPlainSocketImpl.java:188)
    at java.net.SocksSocketImpl.connect(SocksSocketImpl.java:392)
    at java.net.Socket.connect(Socket.java:607)
    at org.apache.http.conn.socket.PlainConnectionSocketFactory.connectSocket(PlainConnectionSocketFactory.java:75)
    at org.apache.http.impl.conn.DefaultHttpClientConnectionOperator.connect(DefaultHttpClientConnectionOperator.java:142)
    ... 29 more
    
```

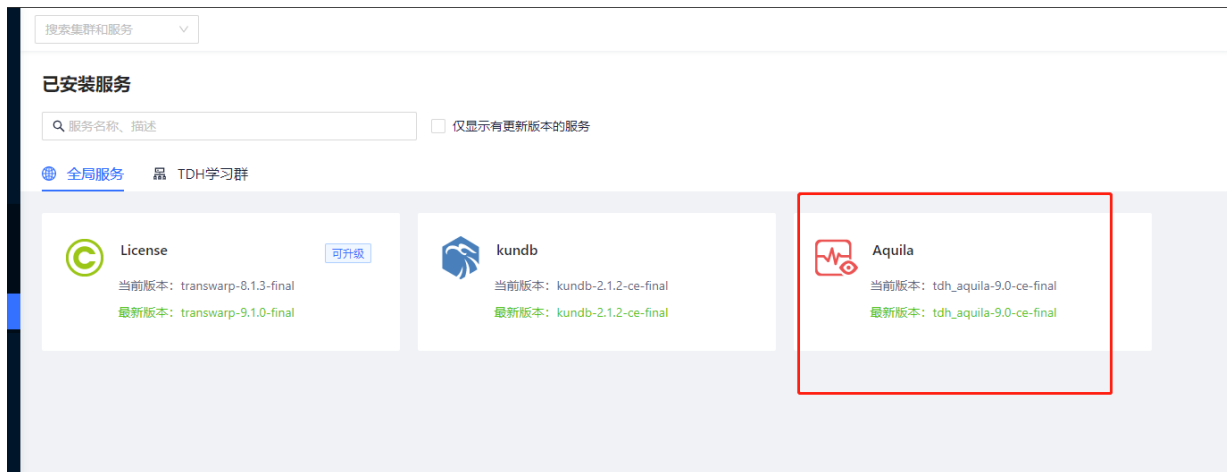
【核心原因】

未安装Aquila

【解决方法】

检查Aquila里面的角色列表是否都是Running，如果存在则启动它，启动失败，则先停止Aquila后删除，

重新配置



8.3.6. TOS相关问题

8.3.6.1. TOS Registry启动失败，报错显示No enough TOS_REGISTRY are healthy...has no running container

【报错信息】

```
2023-06-01T14:52:44.911 [Master] Waiting TOS Registry (TOS,yq-li-centos7-23-37)
to become Healthy within 600 s
2023-06-01T14:54:44.711 [Master] Task 2064 timed out after 120000ms.
2023-06-01T14:54:44.715 [Master] The Task 2064 run failed:
java.util.concurrent.CancellationException
    at java.util.concurrent.FutureTask.report(FutureTask.java:121)
    at java.util.concurrent.FutureTask.get(FutureTask.java:192)
    at
org.springframework.util.concurrent.ListenableFutureTask.done(ListenableFutureTask.java:83)
    at java.util.concurrent.FutureTask.finishCompletion(FutureTask.java:384)
    at java.util.concurrent.FutureTask.cancel(FutureTask.java:180)
    at
io.transwarp.manager.master.operation.execution.TaskDriver.lambda$submitTask$0(TaskDriver.java:75)
```

kube-scheduler is running as 6708ad7088b00d5e66f842f2bdea22b48d6bc5e0dbd3af83c1861f2986eda0a0

TOS_MASTER is running on yq-li-centos7-23-33

etcd --data-d is running as da5060b213a0497d17dd6f84ce3357b8bc97af23e2e36c0212264131ba4df49a

kube-apiserver is running as 0a93a299416fd364b8e43885e1d41bdd9a8038d5fc958e05ec976ca0ac6cddc3

kube-controller is running as b5f6e6eb74aaafd58dd04a90c0135ac87bc795c92c653913aa5f92d91fd327b4

kube-scheduler is running as 5e9c0e8517c311c42d70a74c958ed6fbbc933a937b64096b288da90077a7acd2

No enough TOS_REGISTRY are healthy

TOS_REGISTRY is not running on yq-li-centos7-23-37

/entrypoint.sh has no running container

【解决方法】

在有问题的节点执行 `docker load -i /etc/tos/conf/tos.tar.gz`，再重试即可。

8.3.7. Aquila相关问题

8.3.7.1. DBA产品页面上显示的stage倾斜值在什么范围比较正常

【用户问题】

这个倾斜的值在什么范围比较正常，版本是tdh-aquila-9.0-ce

RGANIZATION = T0.INST_ID and T0.BZ= '2.currency' WHERE T0.DATA_DATE='2023-07-01' GROUP BY T0.INST_ID,T0.BZ,T0.WD_O,T0.WD_D,T0.WD_H,T0.WD_K,T0.WD_L,T0.WD_M,T0.WD_E,T0.WD_F,T0.WD_B,T0.WD_G,T0.WD_J,T0.WD_I,T0.WD_C,T0.WD_N,T0.WD_P,T0.WD_A apply at HierarchicalPerf.scale:230

< 1 > 10 / page

Stage ID	Description	Submission Time	Completed Time	Duration	Shuffle Read	Shuffle Write	Task Skewness	Tasks
5	Stage 5	07/21 09:35:23	07/21 09:38:16	2.9 m	2.1GB	0	1.63	0 / 10 / 0 / 0
7	Stage 7	07/21 09:22:57	07/21 09:26:47	3.8 m	0	1.8GB	77.13	0 / 34 / 0 / 0
9	Stage 9	07/21 09:26:47	07/21 09:30:21	3.6 m	1.3GB	1.8GB	1.15	0 / 33 / 0 / 0
10	Stage 10	07/21 09:30:21	07/21 09:31:42	1.4 m	1.5GB	882.3MB	1.3	0 / 33 / 0 / 0
11	Stage 11	07/21 09:31:42	07/21 09:35:23	3.7 m	660MB	2.8GB	1.14	0 / 19 / 0 / 0

< 1 > 10 / page

【问题答复】

Task Skewness = Stage MaxTaskCost / Stage MinTaskCost 超过100可能就需要关注一下了

8.3.8. 其他问题

8.3.8.1. 如何启停集群，目磁盘不够用了，打算做部分数据的迁移工作

【用户问题】

考虑停止整个集群，迁移数据，如何启停？

【问题答复】

您好，需要先确认磁盘中目标迁移数据是否为hdfs数据，如果磁盘满了的话优先清理日志，把空间释放出来可以在Manager界面做集群停止，迁移数据。Manager界面即可一键启停，右上角有相关按钮。

8.3.8.2. k8s集群的bearer_token路径在哪里，一直找不到

【问题答复】

CCE TOKEN 获取

```
TOKEN=$(cat /var/run/secrets/kubernetes.io/serviceaccount/token)
```

8.3.8.3. 内存不够用经常溢出，日志有个名为executor_heapDump.hprof文件很大，可以删除吗

【用户问题】

其他的都是几M，这个有26G。

【问题答复】

可以删，说明之前crash过，Executor有发生OOM之类情况会自动输出heapdump 文件。

8.4. 端口冲突相关问题

8.4.1. 什么是端口？

所谓的端口，就好像是门牌号一样，客户端可以通过ip地址找到对应的服务器端，但是服务器端是有很多端口的，每个应用程序对应一个端口号，通过类似门牌号的端口号，客户端才能真正地访问到该服务器。为了对端口进行区分，将每个端口进行了编号，这就是端口号。

8.4.2. 如何避免出现端口冲突，保障集群顺利安装以及稳定运行

方法1:

基于集群规划进行资源分配，按照业务切等等。因此如果要安装在同一节点上时，可以在安装的【服务配置】步骤中对默认端口号进行更改，避免重复安装服务导致的默认端口冲突的问题。但如果 Quark 安装在不同节点的话，比如 quark1 在 IDC1，quark2 在 IDC2，那两个服务的端口一样也可以。

社区版常用的端口列表：<https://community.transwarp.cn/question/751>

方法2:

在使用某个端口或是更改端口号时，建议用户在Linux系统终端使用netstat命令查询端口的使用情况。

a. 查看某个具体端口的占用情况:

```
netstat -anp | grep <port>
```

b. 查看全部端口占用情况:

```
netstat -anp  
//或  
netstat -tln
```

8.4.3. 怎么判断是出现了端口冲突

出现端口冲突的常见报错信息如下:

```
Failed to bind to: /0.0.0.0:51888 Caused by: java.net.BindException: Address already in use a
```

报错中一般会直接报出冲突的端口号，比如该报错中冲突的端口号为 51888。

页面有时可以显示出具体涉及的端口。如果没有显示，可以通过下述方式确认冲突的端口信息。



- 查看是哪个服务哪个角色启动报错
- 登录manager页面查看这个服务/角色对应有哪些端口配置
- 通过netstat -nalp | grep <目标端口> 查看是否占用

【解决方法】

- a. Manager 对应服务配置中查看该端口号对应信息。
- b. 运行命令查出端口冲突的进程 ID:

```
netstat -anp | grep <step1 中的端口号>
```

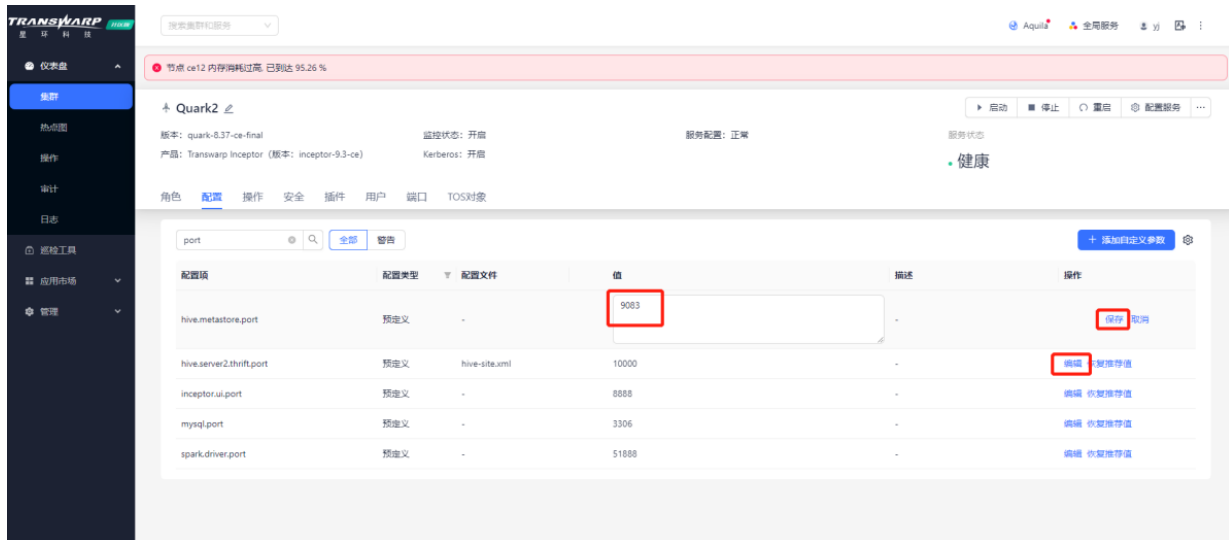
c. 查出具体哪些服务端口冲突:

```
ps -ef | grep <step2 中的进程 ID>
```

8.4.4. 如何解决端口冲突

【解决方法】

对于发生端口冲突导致不可用的服务，确定其冲突的端口参数之后在【服务-配置】界面对端口参数进行修改(一般默认+10)，修改后点击右上角的配置服务，然后重启该服务即可。如下所示：



8.4.5. 其他获取冲突端口信息的方式

当服务安装/启动失败时，可以进入集群中的任意节点，运行kubect1获取pod运行在节点上的信息：

```
kubect1 get po -owide |grep <name>
kubect1 logs -f <有问题的podid>
```

输出结果示例如下，下方示例可以看出该服务的端口[METASTORE_PORT]的默认端口号9093发生了冲突：

```
mysql: [Warning] Using a password on the command line interface can be insecure.
+ return 0
++ grep 'data-dictionary\b' /tmp/tmp.hRA8UHvmzd
++ wc -l
+ '[' 1 == 0 ']'
+ return 0
+ set -e
+ checkPort METASTORE_PORT 9093
+ set +x
ERROR: port [9093] for [METASTORE_PORT] has been bound!
```

8.4.6. 端口连通性测试

- SSH测试：SSH最常用的功能之一就是实现远程登录主机，可以用指定端口号登录的方式来测试端口连通性。测试过程中出现“Connection established”关键字表示端口通，出现“Connection refused”关键字则不通。此方法适用于TCP协议

```
ssh -v -p <port> <ip>
```

- Telnet测试: telnet命令可以进行远程登录计算机,也可以来测试端口的连通性。测试过程中出现“Connection refused”关键字则不通。此方法适用于TCP协议

```
telnet <ip> <port>
```

- Wget测试: wget是一个下载文件的工具,支持http, https, ftp协议,可用于测试端口连通性。

```
wget <ip:port>
```

- NC (Netcat) 测试: nc是一个功能强大的网络工具,对于端口连通性测试来说支持TCP和UDP协议端口测试,也可以对指定范围内的端口进行批量扫描。

```
nc -vuz <ip> <port>
```

8.4.7. 一些排错示例

8.4.7.1. 示例一. 通过TOS启动yarn-nodemanager一个节点启动失败, pod无法启动, 状态显示CCrashLoopBackOff

【报错信息】

运行kubectl命令后获取pod运行在节点上的信息, 通过hadoop-yarn-nodemanager的日志发现是端口被kube-apiserv服务占用。

```

75 2024-01-10T16:43:07.158380 [Master] ===== Task 329 start to run. =====
76 2024-01-10T16:43:07.159607 [Master] Starting task local part ...
77 2024-01-10T16:43:07.160554 [Master] Start handle kubectl task
78 2024-01-10T16:43:07.162519 [Master] Target yaml path on manager is /var/lib/transwarp-manager/master/content/resources/services/yarn/hadoop-yarn-nodemanager.yaml
79 2024-01-10T16:43:07.162543 [Master] Start to generate hadoop-yarn-nodemanager.yaml on manager...
80 2024-01-10T16:43:07.243676 [Master] generated hadoop-yarn-nodemanager.yaml on [Manager]
81 2024-01-10T16:43:07.425699 [Master] Start [Create Roles] ...
82 2024-01-10T16:43:07.425737 [Master] Start executing [kubectl --server=https://127.0.0.1:6443 --certificate-authority=/srv/kubernetes/ca.pem --client-certificate=/srv/kubernetes/admin.pem --client-key=/srv/kubernetes/admin-key.pem scale --replicas=3 -f /var/lib/transwarp-manager/master/content/resources/services/yarn/hadoop-yarn-nodemanager.yaml]
83 2024-01-10T16:43:07.379123 [Master] [Create Roles] success.
84 2024-01-10T16:43:07.379179 [Master] Task local part ended.
85 2024-01-10T16:43:07.579195 [Master] Starting task remote part ...
86 2024-01-10T16:43:07.581273 [Master] Waiting 3 YARN_NODEMANAGER role(s) in YARNM to become Healthy within 600 s
87 2024-01-10T16:53:07.586509 [Master] Latest health check result of roles:
88 DAEMON_CHECK_HEALTHY at 2024-01-10T16:53:05.614
89 YARN_NODEMANAGER on noah-04.bigdata.bytc.qianxin-inc.cn has pod hadoop-yarn-nodemanager-yarn1-6f99585f47-59asc with status RunningVITAL_SIGN_CHECK_HEALTHY at 2024-01-10T16:53:00.655
90 Connection check result of NodeManager noah-04.bigdata.bytc.qianxin-inc.cn is HEALTHYDAEMON_CHECK_WARNING at 2024-01-10T16:53:05.614
91 YARN_NODEMANAGER on noah41.bigdata.bytc.qianxin-inc.cn has pod hadoop-yarn-nodemanager-yarn1-6f99585f47-v7gja with status CrashLoopBackOffVITAL_SIGN_CHECK_DOWN at 2024-01-10T16:53:00.655
92 Connection check result of NodeManager noah41.bigdata.bytc.qianxin-inc.cn is DOWNDAEMON_CHECK_HEALTHY at 2024-01-10T16:53:05.614
93 YARN_NODEMANAGER on noah222.bigdata.bytc.qianxin-inc.cn has pod hadoop-yarn-nodemanager-yarn1-6f99585f47-6eccc with status RunningVITAL_SIGN_CHECK_HEALTHY at 2024-01-10T16:53:00.655
94 Connection check result of NodeManager noah222.bigdata.bytc.qianxin-inc.cn is HEALTHY
95 2024-01-10T16:53:07.586744 [Master] Fail to run task remote part: io.transwarp.manager.master.operation.execution.entity.TaskWarningException: java.lang.IllegalStateException: 3 YARN_NODEMANAGER role(s) in YARNM didn't become healthy within 180 s
96   at io.transwarp.manager.master.operation.execution.localrunner.RubectlTaskLocalRunner.postRemote(RubectlTaskLocalRunner.java:896)
97   at io.transwarp.manager.master.operation.execution.TaskDriverTaskRunner.runTask(TaskDriver.java:287)
98   at io.transwarp.manager.master.operation.execution.TaskDriverTaskRunnerHelper$FastClassBySpringAOPWapper.invoke(<generated>)
99   at org.springframework.cglib.proxy.MethodProxy.invoke(MethodProxy.java:218)
100  at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.invokeJoinpoint(CglibAopProxy.java:793)
101  at org.springframework.aop.framework.ReflectiveMethodInvocation.proceed(ReflectiveMethodInvocation.java:163)
102  at org.springframework.aop.framework.CglibAopProxy$CglibMethodInvocation.proceed(CglibAopProxy.java:763)
103  at org.springframework.aop.interceptor.AsyncExecutionInterceptor.lambda$invoke$0(AsyncExecutionInterceptor.java:115)
104  at java.base/java.util.concurrent.FutureTask.run(FutureTask.java:264)
105  at java.base/java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1128)
106  at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:628)
107  at java.base/java.lang.Thread.run(Thread.java:829)
108 Caused by: java.lang.IllegalStateException: 3 YARN_NODEMANAGER role(s) in YARNM didn't become healthy within 180 s
109   at io.transwarp.manager.master.operation.execution.localrunner.AbstractTaskLocalRunner.waitForHealthy(AbstractTaskLocalRunner.java:721)
110   at io.transwarp.manager.master.operation.execution.localrunner.RubectlTaskLocalRunner.postRemote(RubectlTaskLocalRunner.java:887)
111   ... 11 more
112 Caused by: org.awaitility.core.ConditionTimeoutException: Condition with lambda expression in io.transwarp.manager.master.operation.execution.localrunner.AbstractTaskLocalRunner was not fulfilled within 10 minutes.
113   at org.awaitility.core.ConditionAwaiter.await(ConditionAwaiter.java:167)

```

```

HostPathType: /var/lib/docker/volumes
default-token-kvdt9t:
  Type: Secret (a volume populated by a Secret)
  secretName: default-token-kvdt9t
  Optional: false
OS Class: hadEffort
Node-Selectors:
  node.kubernetes.io/nodename: !hostname
Tolerations:
  node.kubernetes.io/not-ready: NoExecute for 1800s
  node.kubernetes.io/unreachable: NoExecute for 1800s
Events:
Type Reason Age From Message
Normal Scheduled <unknown> default-scheduler Successfully assigned default/hadoop-yarn-nodemanager-yarn1-6f9985f47-1zzrn to noah44.bigdata.bjyc.qianxin-inc.cn
Normal Pulling 14s (x4 over 66s) kubelet, noah44.bigdata.bjyc.qianxin-inc.cn Pulling image "transwarp/yarn:transwarp-9.3-ce-final"
Normal Pulled 14s (x4 over 66s) kubelet, noah44.bigdata.bjyc.qianxin-inc.cn Successfully pulled image "transwarp/yarn:transwarp-9.3-ce-final"
Normal Created 14s (x4 over 66s) kubelet, noah44.bigdata.bjyc.qianxin-inc.cn Created container hadoop-yarn-nodemanager-yarn1
Normal Started 14s (x4 over 66s) kubelet, noah44.bigdata.bjyc.qianxin-inc.cn Started container hadoop-yarn-nodemanager-yarn1
Warning Backoff 11s (x4 over 58s) kubelet, noah44.bigdata.bjyc.qianxin-inc.cn Back-off restarting failed container

[root@noah44 var]# kubectl get pod
NAME READY STATUS RESTARTS AGE
alertmanager-aquila-5c745b4cfb-dtxq2 1/1 Running 0 66m
aquila-agent-aquila-5706cbb96-vrt42 1/1 Running 0 66m
aquila-agent-aquila-5706cbb96-x81tm 1/1 Running 0 66m
aquila-agent-aquila-5706cbb96-zdb1h 1/1 Running 0 66m
aquila-server-aquila-83f9584f-1jpl6 1/1 Running 0 65m
dba-service-aquila-666cd4889-11f45 1/1 Running 0 61m
hadoop-hdfs-datanode-hdfs1-6b6e77fdcc-4dfnn 1/1 Running 0 54m
hadoop-hdfs-datanode-hdfs1-6b6e77fdcc-8w57d 1/1 Running 0 54m
hadoop-hdfs-datanode-hdfs1-6b6e77fdcc-g9lmb 1/1 Running 0 54m
hadoop-hdfs-ncdfs-hdfs1-74dbd946-1h96 1/1 Running 0 55m
hadoop-hdfs-journalnode-hdfs1-584d9c5586-5twtw 1/1 Running 0 55m
hadoop-hdfs-journalnode-hdfs1-584d9c5586-fygsd 1/1 Running 0 55m
hadoop-hdfs-journalnode-hdfs1-584d9c5586-r8zjg 1/1 Running 0 55m
hadoop-hdfs-namenode-hdfs1-56f974db8c-34n8n 2/2 Running 0 55m
hadoop-hdfs-namenode-hdfs1-56f974db8c-rlkhn 2/2 Running 0 55m
hadoop-yarn-historyserver-yarn1-64b6d9f64d-86cnd 1/1 Running 0 34m
hadoop-yarn-nodemanager-yarn1-6f9985f47-1j2h 1/1 Running 0 76s
hadoop-yarn-nodemanager-yarn1-6f9985f47-1x55b 1/1 Running 0 76s
hadoop-yarn-nodemanager-yarn1-6f9985f47-1zzrn 0/1 CrashLoopBackOff 3 76s
hadoop-yarn-resourcemanager-yarn1-6957cfcc45-h2vc7 1/1 Running 0 45m
hadoop-yarn-resourcemanager-yarn1-6957cfcc45-qj7n6 1/1 Running 0 45m
hadoop-yarn-cimelineserver-yarn1-6c49f5645-jvdhj 1/1 Running 0 62m
insights-server-aquila-8699ff9f8-2r67 1/1 Running 0 68m
kube-state-metrics-aquila-f876fffb5-92d7k 1/1 Running 0 68m
kubernetes-external-kubernetes-3cbdc47cdd-6v1ns 1/1 Running 0 72m

```

```

2024-01-11 19:23:12.639 INFO org.apache.hadoop.mapred.indexcache: IndexCache created with max memory = 10485760
2024-01-11 19:23:12.662 INFO org.apache.hadoop.service.AbstractService: Service mapreduce_shuffle failed in state STARTED
org.jboss.netty.channel.ChannelException: Failed to bind to: 0.0.0.0/0.0.0.0:13562
    at org.jboss.netty.bootstrap.ServerBootstrap.bind(ServerBootstrap.java:272)
    at org.apache.hadoop.mapred.ShuffleHandler.serviceStart(ShuffleHandler.java:344)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:194)
    at org.apache.hadoop.yarn.server.nodemanager.containermanager.AuxServices.serviceStart(AuxServices.java:313)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:194)
    at org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl.serviceStart(ContainerManagerImpl.java:668)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:194)
    at org.apache.hadoop.service.composite.service.serviceStart(CompositeService.java:122)
    at org.apache.hadoop.yarn.server.nodemanager.NodeManager.initAndStartNodeManager(NodeManager.java:937)
    at org.apache.hadoop.yarn.server.nodemanager.NodeManager.main(NodeManager.java:1016)
Caused by: java.net.BindException: Address already in use
    at sun.nio.ch.Net.bind0(Native Method)
    at sun.nio.ch.Net.bind(Net.java:436)
    at sun.nio.ch.ServerSocketChannelImpl.bind(ServerSocketChannelImpl.java:214)
    at sun.nio.ch.ServerSocketAdaptor.bind(ServerSocketAdaptor.java:74)
    at org.jboss.netty.channel.socket.nio.NioServerBoss.registerTask.run(NioServerBoss.java:193)
    at org.jboss.netty.channel.socket.nio.AbstractNioSelector.processTaskQueue(AbstractNioSelector.java:391)
    at org.jboss.netty.channel.socket.nio.AbstractNioSelector.run(AbstractNioSelector.java:315)
    at org.jboss.netty.channel.socket.nio.NioServerBoss.run(NioServerBoss.java:42)
    at org.jboss.netty.util.ThreadRenamingRunnable.run(ThreadRenamingRunnable.java:108)
    at org.jboss.netty.util.internal.DeadLockProofWorker$1.run(DeadLockProofWorker.java:42)
    at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)
    at java.lang.Thread.run(Thread.java:745)
2024-01-11 19:23:12.684 INFO org.apache.hadoop.service.AbstractService: Service org.apache.hadoop.yarn.server.nodemanager.containermanager.AuxServices failed in state STARTED
org.jboss.netty.channel.ChannelException: Failed to bind to: 0.0.0.0/0.0.0.0:13562
    at org.jboss.netty.bootstrap.ServerBootstrap.bind(ServerBootstrap.java:272)
    at org.apache.hadoop.mapred.ShuffleHandler.serviceStart(ShuffleHandler.java:344)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:194)
    at org.apache.hadoop.yarn.server.nodemanager.containermanager.AuxServices.serviceStart(AuxServices.java:313)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:194)
    at org.apache.hadoop.yarn.server.nodemanager.containermanager.ContainerManagerImpl.serviceStart(ContainerManagerImpl.java:668)

```

【解决方法】

参考前述章节<怎么判断是出现了端口冲突>以及<其他获取冲突端口信息的方式>章节获取服务端端口信息，按照<如何解决端口冲突>章节内容修改对应端口即可。

8.4.7.2. 示例二. 初始化HDFS失败，启动NameNode报错

【报错信息】

状态	描述	开始时间	结束时间	操作
SUCCESSFUL	生成依赖的服务的配置	2023-07-10 19:03:15	2023-07-10 19:03:15	查看日志
SUCCESSFUL	生成 Collector 配置	2023-07-10 19:03:15	2023-07-10 19:03:15	查看日志
SUCCESSFUL	配置Name Node(test1)	2023-07-10 19:03:15	2023-07-10 19:03:16	查看日志
SUCCESSFUL	配置Https(test1)	2023-07-10 19:03:16	2023-07-10 19:03:16	查看日志
SUCCESSFUL	配置Data Node(test1)	2023-07-10 19:03:16	2023-07-10 19:03:17	查看日志
SUCCESSFUL	通知 Aquila 重载配置文件	2023-07-10 19:03:17	2023-07-10 19:03:18	查看日志
FAILED	启动Name Node(test1)的角色	2023-07-10 19:03:18	2023-07-10 19:06:18	重试 查看日志
WAITING	启动Https(test1)的角色			查看日志
WAITING	启动Data Node(test1)的角色			查看日志

【核心原因】

查看日志后显示端口被占用


```

Caused by: java.io.IOException: Failed to bind to test1/172.16.48.161:50070
    at org.eclipse.jetty.server.ServerConnector.openAcceptChannel(ServerConnector.java:346)
    at org.eclipse.jetty.server.ServerConnector.open(ServerConnector.java:307)
    at org.apache.hadoop.http.HttpServer2.bindListener(HttpServer2.java:1301)
    at org.apache.hadoop.http.HttpServer2.bindForSinglePort(HttpServer2.java:1332)
    ... 9 more
Caused by: java.net.BindException: Address already in use
    at sun.nio.ch.Net.bind0(Native Method)
    at sun.nio.ch.Net.bind(Net.java:461)
    at sun.nio.ch.Net.bind(Net.java:453)
    at sun.nio.ch.ServerSocketChannellImpl.bind(ServerSocketChannellImpl.java:222)
    at sun.nio.ch.ServerSocketAdaptor.bind(ServerSocketAdaptor.java:85)
    at org.eclipse.jetty.server.ServerConnector.openAcceptChannel(ServerConnector.java:342)
    ... 12 more
2023-07-10 11:03:27.478 INFO org.apache.hadoop.metrics2.impl.MetricsSystemImpl: Stopping NameNode metrics system...
2023-07-10 11:03:27.478 INFO org.apache.hadoop.metrics2.impl.MetricsSystemImpl: NameNode metrics system stopped.
2023-07-10 11:03:27.479 INFO org.apache.hadoop.metrics2.impl.MetricsSystemImpl: NameNode metrics system shutdown complete.
2023-07-10 11:03:27.479 ERROR org.apache.hadoop.hdfs.server.namenode.NameNode: Failed to start namenode.
java.net.BindException: Port in use: test1:50070
    at org.apache.hadoop.http.HttpServer2.constructBindException(HttpServer2.java:1314)
    at org.apache.hadoop.http.HttpServer2.bindForSinglePort(HttpServer2.java:1336)
    at org.apache.hadoop.http.HttpServer2.openListeners(HttpServer2.java:1395)
    at org.apache.hadoop.http.HttpServer2.start(HttpServer2.java:1250)
    at org.apache.hadoop.hdfs.server.namenode.NameNodeHttpServer.start(NameNodeHttpServer.java:281)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.startHttpServer(NameNode.java:898)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.initialize(NameNode.java:720)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.<init>(NameNode.java:966)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.<init>(NameNode.java:939)
    at org.apache.hadoop.hdfs.server.namenode.NameNode.createNameNode(NameNode.java:1710)

```

【解决方法】

参考前述章节<怎么判断是出现了端口冲突>以及<其他获取冲突端口信息的方式>章节获取服务端口信息，按照<如何解决端口冲突>章节内容修改对应端口即可。

8.4.7.3. 示例三. 启动Quark失败，查看/var/log下的server日志，报错显示“Failed to bind SparkUI, java.net.BindException:Address already in use (Bind failed): Service ‘SparkUI’ failed after 0 retries!”

【报错信息】

```

2023-04-10 18:55:57,095 INFO restful.SparkRestful: [main()] - CoreEngine initially load restful api class
2023-04-10 18:56:56,914 INFO leviathan.TimedEventTracker: [Leviathan Monitor Timer()] - [Leviathan][0]JVM
2023-04-10 18:56:57,310 ERROR ui.SparkUI: [main()] - Failed to bind SparkUI
java.net.BindException: Address already in use (Bind failed): Service 'SparkUI' failed after 0 retries!
    at java.net.PlainSocketImpl.socketBind(Native Method)
    at java.net.AbstractPlainSocketImpl.bind(AbstractPlainSocketImpl.java:387)

```

【核心原因】

可以参考该链接<<https://community.transwarp.cn/question/751>>文末的端口清单，可以看到是8888端口存在冲突情况，被其他进程占用了。

【解决方法】

Manager访问服务配置列表，修改端口号之后点击右上角配置服务然后重启服务即可。

8.5. 客户端相关问题

TIPS:



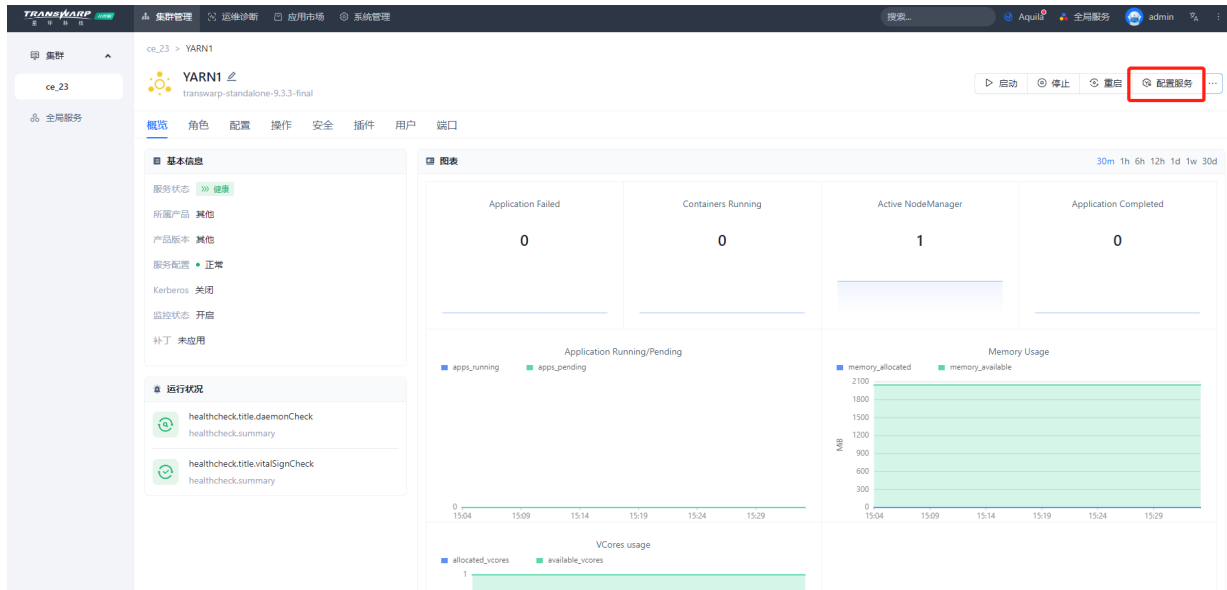
- TDH社区开发版（单机版）的Manager中，“随产品包上传”以及“Manaegr自带”均可以下载使用TDH客户端。
- TDH社区版（分布式）中，请选择“随产品包上传”。注意：请不要选择“Manager自带”哦。

社区开发版与社区版针对客户端部署有不同的方式，具体请参考各自的手册

8.5.1. 点击下载客户端无反应

【解决方法】

可能是重启过容器或其他操作，重新配置下Yarn服务再进行下载即可



8.5.2. Manager中TDH-Client下载失败

【问题概述】

在TDH社区开发版中，下载TDH-Client时，点击两种下载方式后，只跳转页面，但并未出现下载弹窗，因此无法完成下载客户端压缩包的操作。

【初步排错】

查看/var/lib/transwarp-manager/master/content/resources/bundled_tdh_client/目录下是否有tdh-client.tar文件，且文件大小在变化。（正在新建压缩包）

```
[root@ce113 resources]# ls
bundled_tdh_client  nodes  tdh_client
```

查看后，并没有上述文件。其中的tdh_client目录是用来打包客户端的临时目录，说明压缩指令已经执行，但是压缩时间过长。

可能原因：集群所安装的机器io较慢导致打包很慢。若3分钟以上仍无法完成打包操作，则表明并非机器慢，需联系社区版技术支持另外排错。

8.5.3. 如何让TDH客户端初始化失效？

【用户问题】

目前有两个集群，TDH与其他数据库，目前在其他数据库集群里，安装了TDH的客户端并执行过source，现在想让这个失效，直接注释掉环境变量的内容。初始化其他数据库客户端后，仍连接TDH的集群，应如何解决呢？这个是需要重启机器才可以的吗？

【问题答复】

您好，重启一个Session即可。同理，若您在操作TDH，连接Inceptor时，若重启新的session，则需要重新执行初始化客户端。

8.5.4. 更换/新增了服务，之前下载的客户端是否会冲突？

【用户问题】

在同一环境下，若在CentOS上安装了HDFS的客户端，现阶段有一个HDFS和Inceptor的客户端的安装包，那么如果把HDFS和Inceptor的安装包再次安装一遍，会不会报错，前者HDFS和后者HDFS出现冲突呀

【问题答复】

如果只是装了其他类型的服务，可以把原来那个删了，直接用个新的；如果是两个相同类型服务，比如hdfs1和hdfs2，可以分别进行下载。

8.5.5. 开启Guardian后，TDH-Client报错，安装了kerberos client，hdfs也报错

【报错信息】

```
ariprtvmddb02 kerberos]# vi /etc/krb5.conf
ariprtvmddb02 kerberos]# hdfs dfs -ls /
log4j.properties is not found. HADOOP_CONF_DIR may be incomplete.
-31 09:49:16,996 INFO [main] util.KerberosUtil (KerberosUtil.java:getDefaultPrincipalPattern(67)) - Using
-31 09:49:17,601 WARN [main] ipc.Client (Client.java:run(779)) - Exception encountered while connecting to
che.hadoop.security.AccessControlException: Client cannot authenticate via:[TOKEN, KERBEROS]
at org.apache.hadoop.security.SaslRpcClient.selectSaslClient(SaslRpcClient.java:182)
at org.apache.hadoop.security.SaslRpcClient.saslConnect(SaslRpcClient.java:407)
at org.apache.hadoop.ipc.Client$Connection.setupSaslConnection(Client.java:623)
at org.apache.hadoop.ipc.Client$Connection.access$2300(Client.java:414)
at org.apache.hadoop.ipc.Client$Connection$2.run(Client.java:829)
at org.apache.hadoop.ipc.Client$Connection$2.run(Client.java:825)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:2145)
at org.apache.hadoop.ipc.Client$Connection.setupIOstreams(Client.java:825)
at org.apache.hadoop.ipc.Client$Connection.access$3800(Client.java:414)
at org.apache.hadoop.ipc.Client.getConnection(Client.java:1643)
at org.apache.hadoop.ipc.Client.call(Client.java:1459)
at org.apache.hadoop.ipc.Client.call(Client.java:1412)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:233)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:118)
at com.sun.proxy.$Proxy9.getFileInfo(Unknown Source)
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.getFileInfo(ClientNamenodeProtocolT
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:422)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invokeMethod(RetryInvocationHandler.java:165)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invoke(RetryInvocationHandler.java:157)
at org.apache.hadoop.io.retry.RetryInvocationHandler$Call.invokeOnce(RetryInvocationHandler.java:95)
at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:359)
```

【核心原因】

开启guardian后没有重新下载客户端

【解决方法】

需要重新下载才可以包含开启安全的配置。

并且，同样需要注意的是，source init属于session级别的指令，因此每次打开一个新的session均需要重新执行source init。

8.6. Waterdrop连接与使用相关问题

8.6.1. Waterdrop是什么，如何安装？

产品介绍

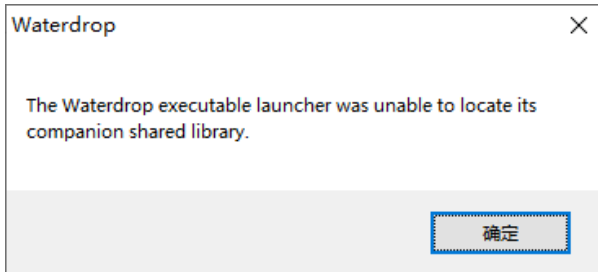
Waterdrop是面向开发者个人的数据库开发IDE，目标是提高开发者的开发效率。产品主要功能有元数据管理、SQL开发、数据对象管理、驱动管理、连接管理、用户管理等等，更多介绍可参考：

<https://community.transwarp.cn/article/931>

链接数据库教程

星环产品支持通过 Beeline、应用程序、外部工具等多种方式连接数据库，Waterdrop的链接教程可参考：<https://community.transwarp.cn/article/368>

8.6.2. Waterdrop下载后，打开报错显示：The Waterdrop executable launcher was unable to locate its companion shared library

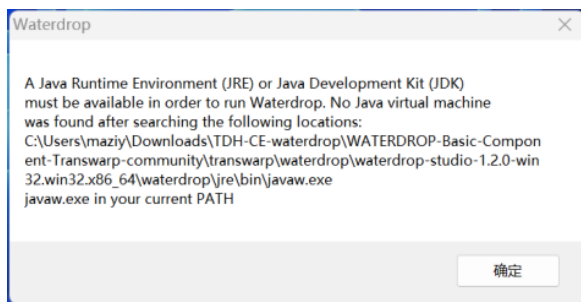


【核心原因】

这个问题一般是安装的jdk版本不对，需更换为jdk1.8。

8.6.3. Windows系统打开Waterdrop时出现无JRE/JDK报错，应如何解决？

【报错情况】



【根因分析】

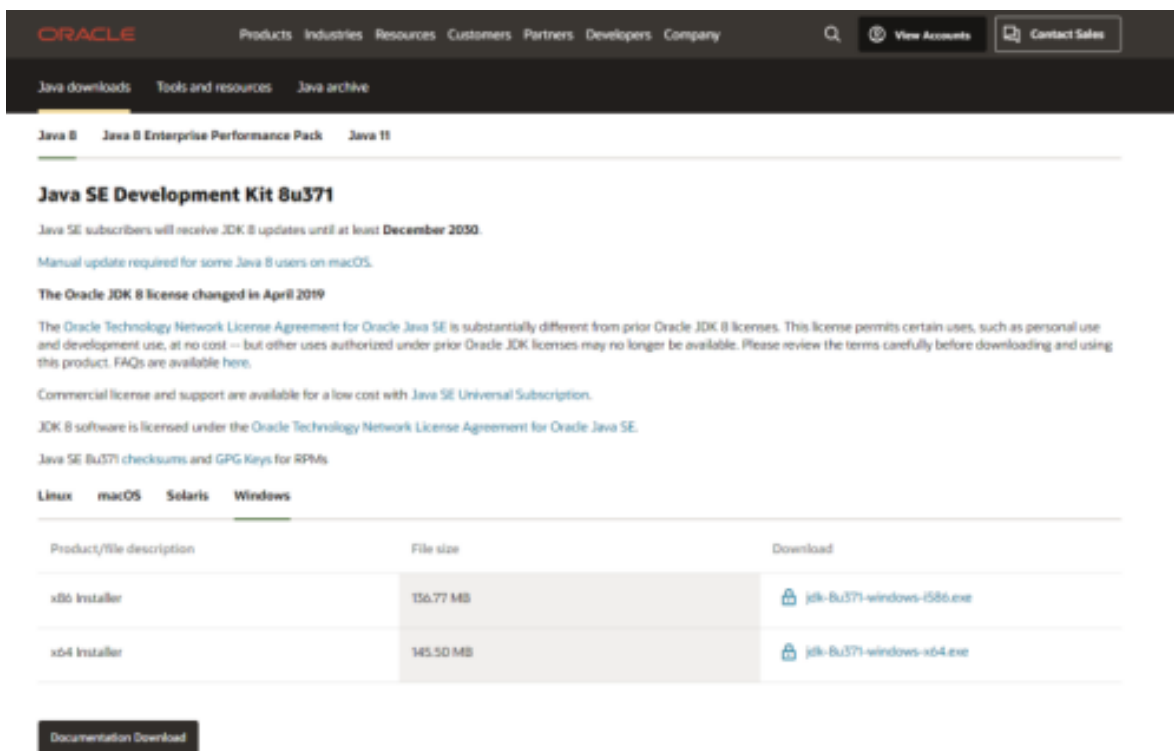
- Waterdrop运行需要JAVA环境，因此需要进行Java环境的配置。
- 环境说明：Windows 11、Java 8

【解决步骤】

1. 第一步：JDK的下载与安装

Jdk下载地址：

<https://www.oracle.com/java/technologies/downloads/#java8> 选择与电脑系统匹配的版本进行下载，Windows x86对应32位系统，Windows x64对应64位系统。MacOS系统下载请参考该文章（<https://community.transwarp.cn/article/detail?id=357>），其他系统可以在官网寻找对应系统即可。



2. 第二步：安装

双击下载的安装包，根据提示完成安装即可。



在此可选择安装的路径，建议使用默认路径。需记录一下安装路径，后续配置环境时会用到。

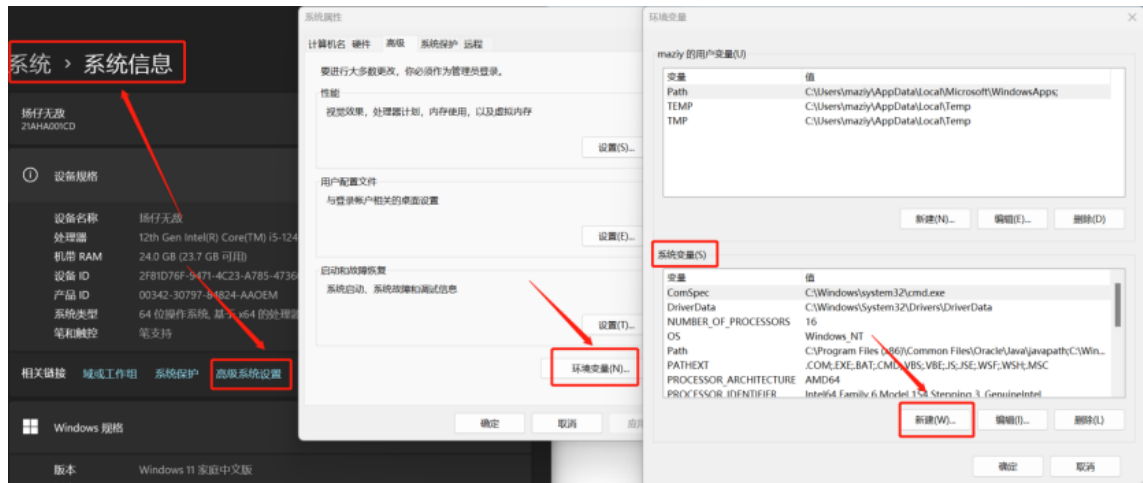


此时，已完成Java的安装。接下来要进行环境变量的配置。

3. 第三步：配置环境变量

a. 打开环境变量

右键我的电脑，点击下面的“属性”按钮，之后在弹窗中选择“高级系统设置”。

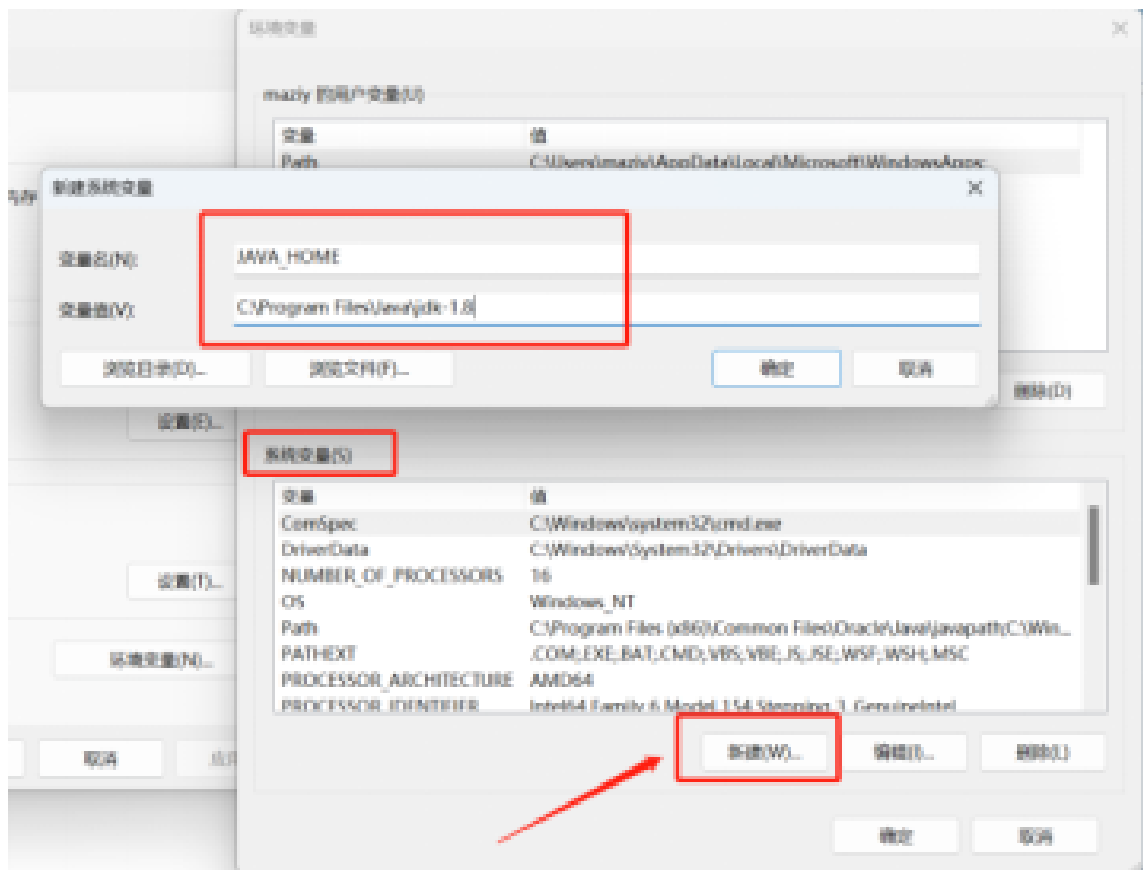


b. 配置JAVA_HOME

新建系统变量

变量名: JAVA_HOME

变量值: 安装路径

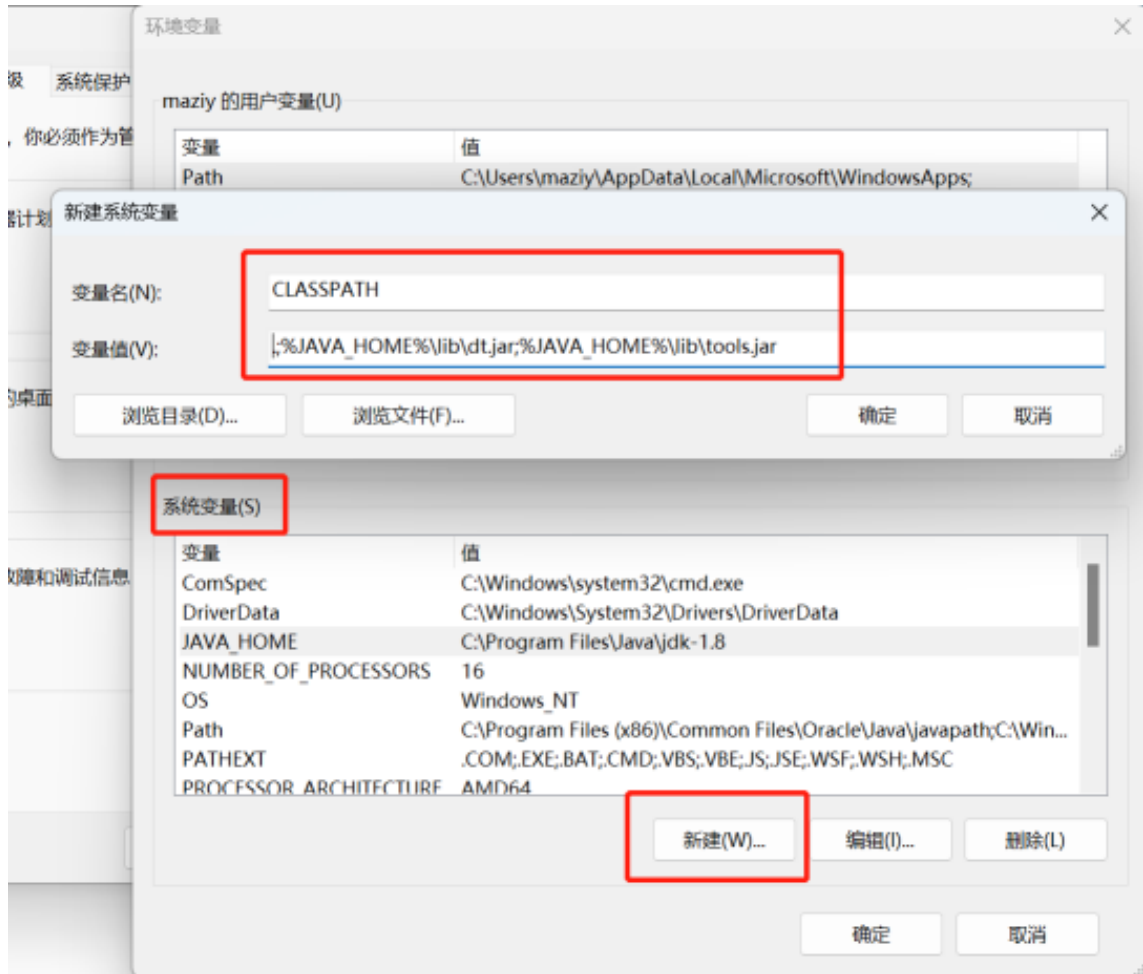


c. 配置ClassPath

新建系统变量

变量名: CLASSPATH

变量值: .;%JAVA_HOME%\lib\dt.jar;%JAVA_HOME%\lib\tools.jar



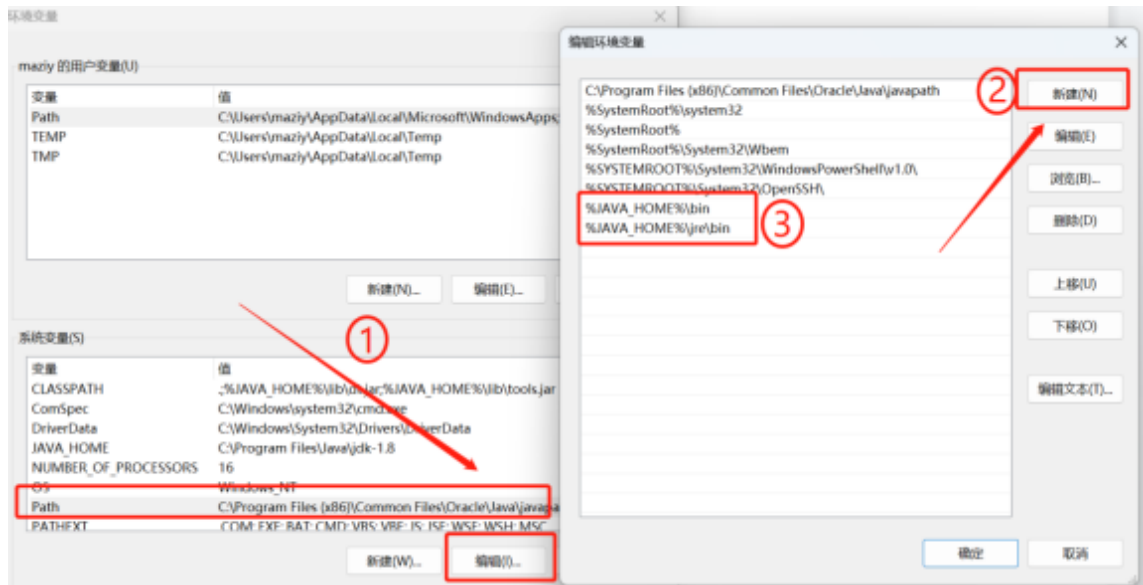
d. 配置PATH

在PATH变量后添加两条变量

编辑环境变量

变量值: %JAVA_HOME%\bin

变量值: %JAVA_HOME%\jre\bin



e. 配置完成后，验证是否安装成功

使用快捷键（win+r），输入“cmd”打开“命令提示符”以后，输入“java -version”。出现下图所示内容，即表示安装成功。

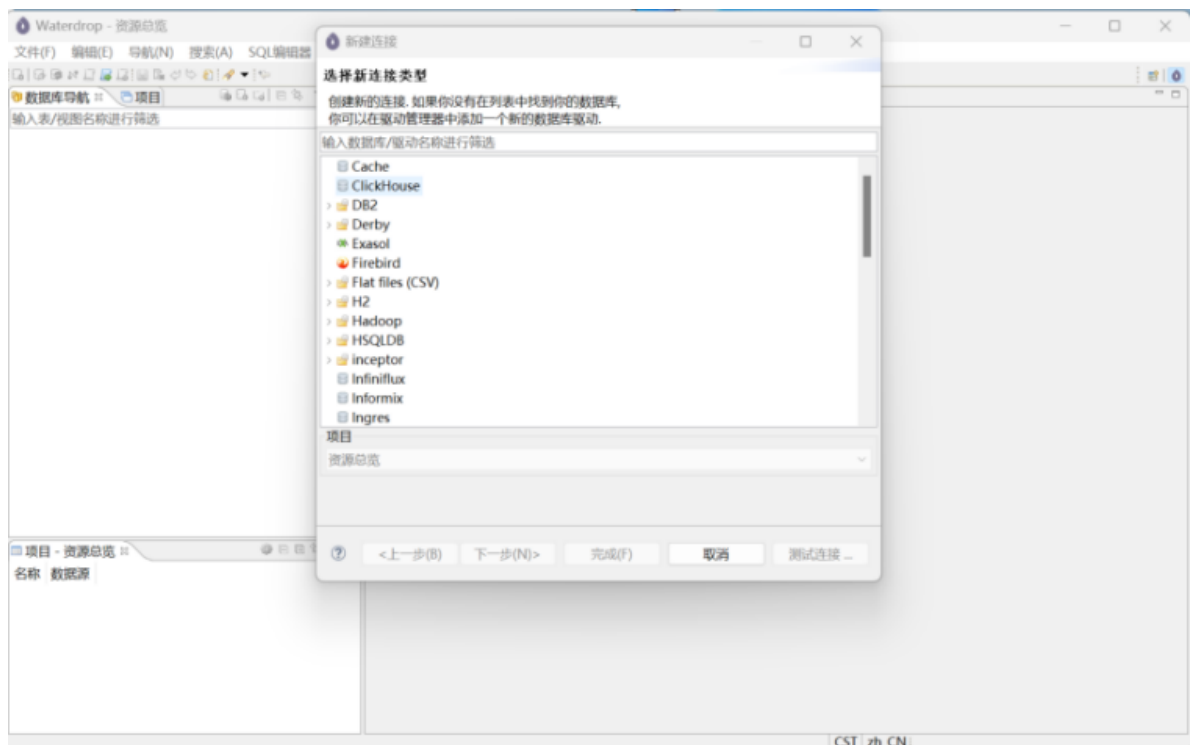
```

C:\Windows\system32\cmd.exe
C:\Users\mazy>java -version
java version "1.8.0_371"
Java(TM) SE Runtime Environment (build 1.8.0_371-b11)
Java HotSpot(TM) 64-Bit Server VM (build 25.371-b11, mixed mode)
C:\Users\mazy>

```

4. 第四步：打开Waterdrop

最后，尝试打开Waterdrop可以顺利打开，问题解决。



8.6.4. Waterdrop 1.2.x 在MacOS Ventura 13.x上无法正常运行

【问题描述】

Waterdrop无法正常运行，解压安装包后尝试双击运行，程序栏出现Waterdrop图标但立即消失，不能成功运行。再次点击后无任何响应，排查解决思路如下。

- 示例环境

Waterdrop 1.2.0

MacOS Ventura 13.4

【解决步骤】

1. **安装JDK环境** 出现该错误，一般与JDK环境有关系。Waterdrop需要依赖于JDK8才能正常工作。若您的机器上尚未安装JDK8，请先安装，安装详情请查看《Mac OS系统如何安装JAVA：<https://community.transwarp.cn/article/357>》文章。
2. **尝试运行后报错** 若您的 MacOS 环境中已经安装了JDK8，但 Waterdrop 依然不能成功运行，打开Waterdrop仍 出现弹窗，参照如下步骤操作。



3. **查询Java路径**

使用 Finder 或终端工具浏览到 /Library/Java/JavaVirtualMachines 并检查那里的 JDK 包名称，记录这边实际看到的 JDK 版本。例如下图中的JDK版本为1.8.0

```
yangzaiwudi@yangzaiwudideMacBook-Pro bin % pwd
/Library/Java/JavaVirtualMachines/jdk1.8.0_202.jdk/Contents/Home/bin
```

4. 右键点击Waterdrop图标，选择“显示包内容”



5. 进入Contents目录，编辑info.plist文件。



进入文档后，向下滚动到文档底部，可以看到下图内容。

6. 编辑内容

此时，需要将上图中红色框中文本移至上方，并修改图中黄色框中内容，修改为步骤3中查询到的实际版本参照如下：

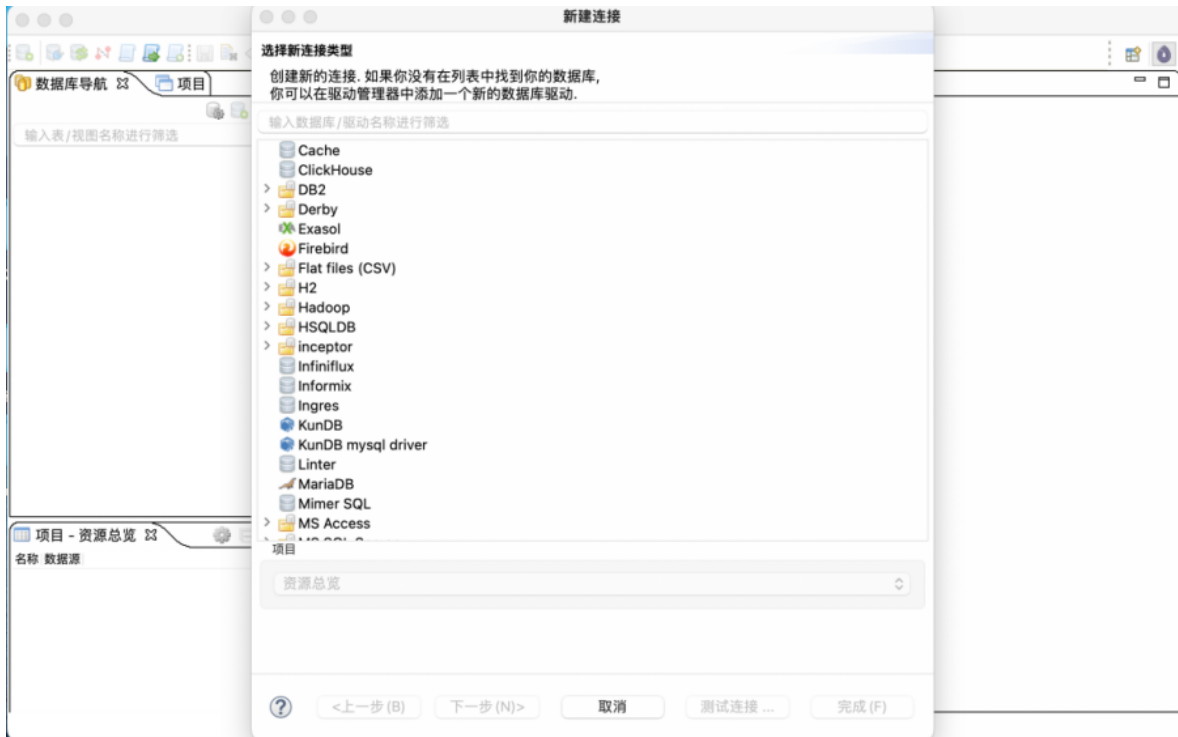
```

</array>
<key>Eclipse</key>
<array>
  <string>-vm</string><string>/Library/Java/JavaVirtualMachines/jdk1.8.0_202.jdk/Contents/Home/bin/java</string>
  <!-- to use a specific Java version (instead of the platform's default) uncomment one of the following options,
       or add a VM found via $/usr/libexec/java_home -V
  <string>-vm</string><string>/System/Library/Java/JavaVirtualMachines/1.6.0.jdk/Contents/Commands/java</string>
-->
  <string>-keyring</string>
</array>
string>~/.eclipse_keyring</string>
</array>
BundleDisplayName</key>
ring>Waterdrop</string>
t>

```

7. 尝试启动

双击Waterdrop图标，确认可以正常使用。



8. 其他信息

如果打开waterdrop出现弹窗：“已损坏，无法打开，你应该将它移到废纸篓。”

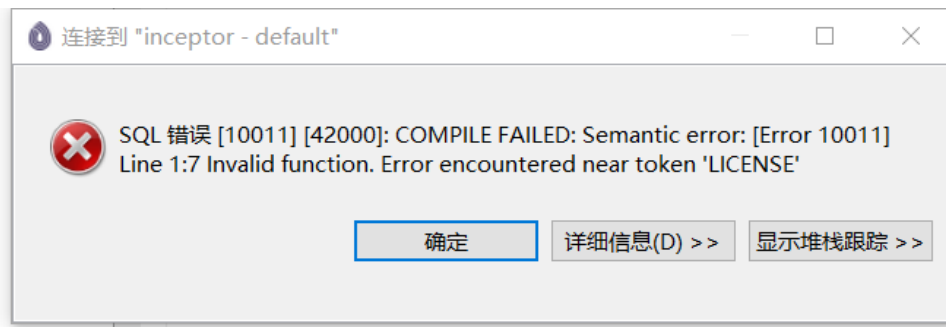
可以先打开 系统偏好设置 → 安全与隐私 → 通用 选项卡，选择 仍然打开；

或在终端中执行以下命令，开启允许任何来源：sudo spctl --master-disable

8.6.5. Waterdrop报错 COMPILER FAILED:Semantic error: [Error 10011] Line 1:7 Invalid function. Error encountered near token ‘LICENSE’

【用户问题】

这个工具是需要申请license吗？

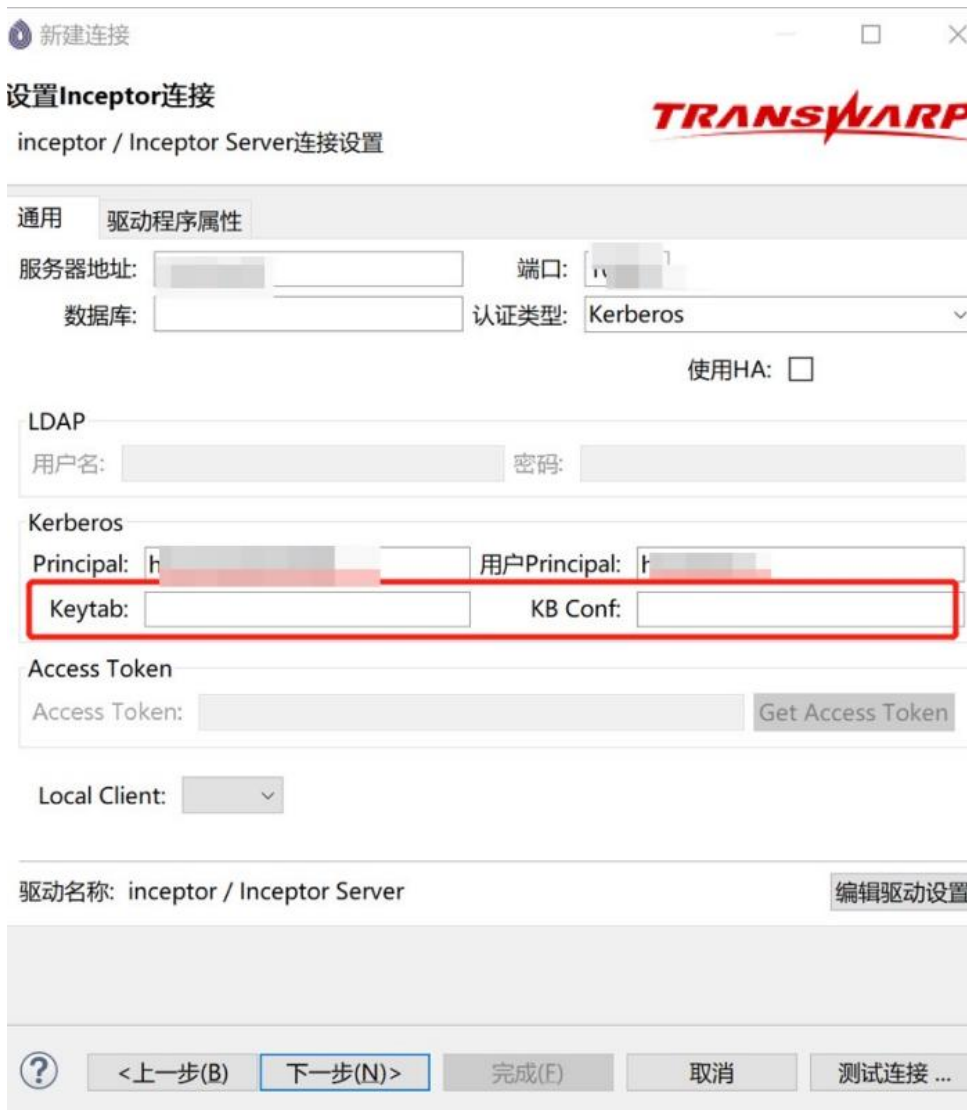


【问题答复】

是下载的版本不对哦，社区版Waterdrop这里做了判断。因此只能对接社区版的Inceptor，不能连企业版inceptor。企业版Inceptor需要下载对应的商业版Waterdrop，可以参考下方手册链接了解如何申请商业版Waterdrop的许可证。 https://docs.transwarp.cn/#/documents-support/docs-detail/document/TDH-DEV_SUITE/7.0/060WaterdropManual?docType=docs%3Fcategory%3D%DH%26index%3D0&docName=Waterdrop%E4%BD%BF%E7%94%A8%E6%89%8B%E5%86%8C

8.6.6. Waterdrop如何获取Keytab文件？

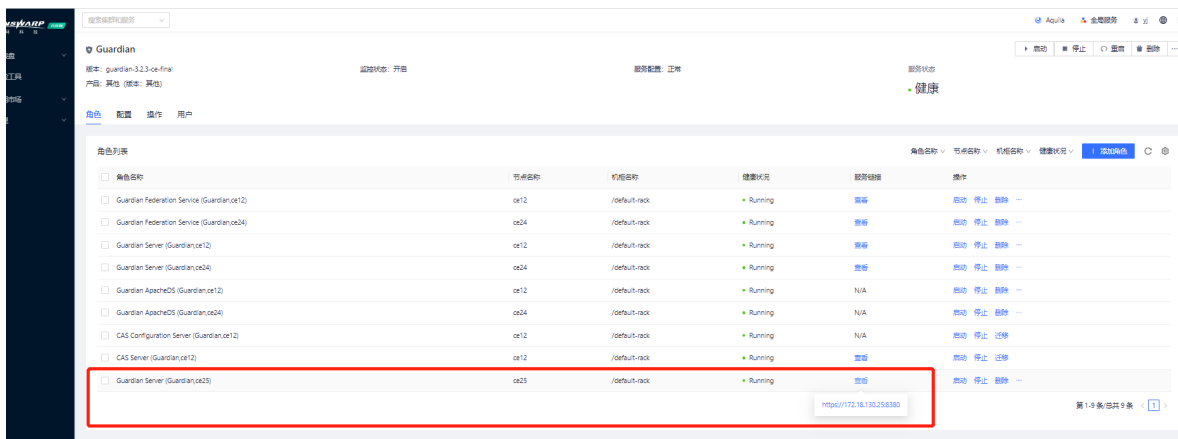
【产品情况】



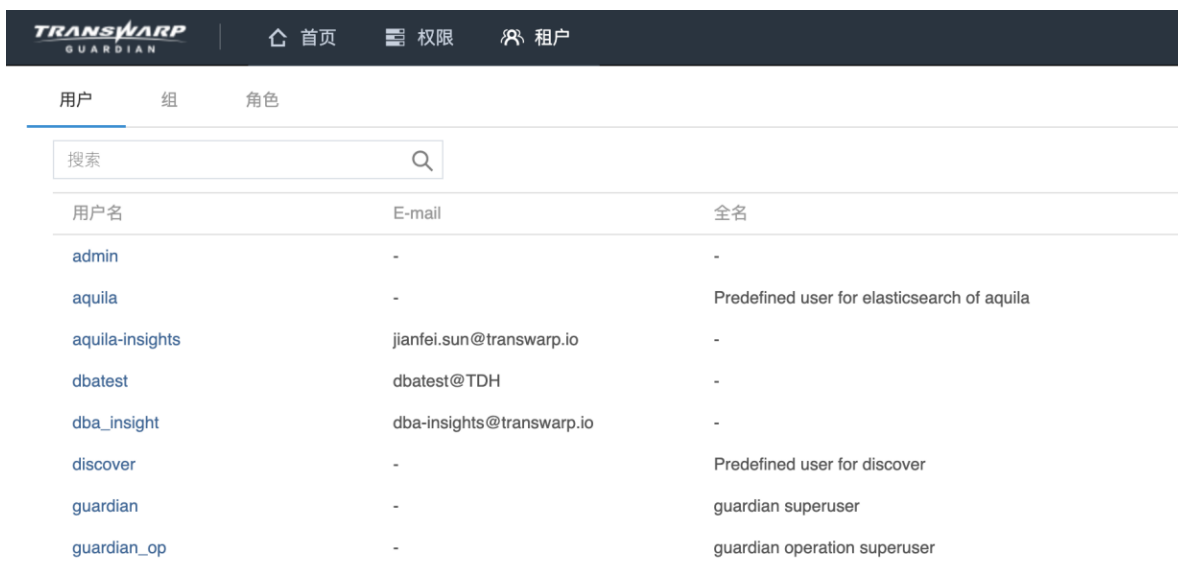
【获取方法】

Keytab文件可以在Guardian中对应租户的详情页获取，具体操作步骤如下：

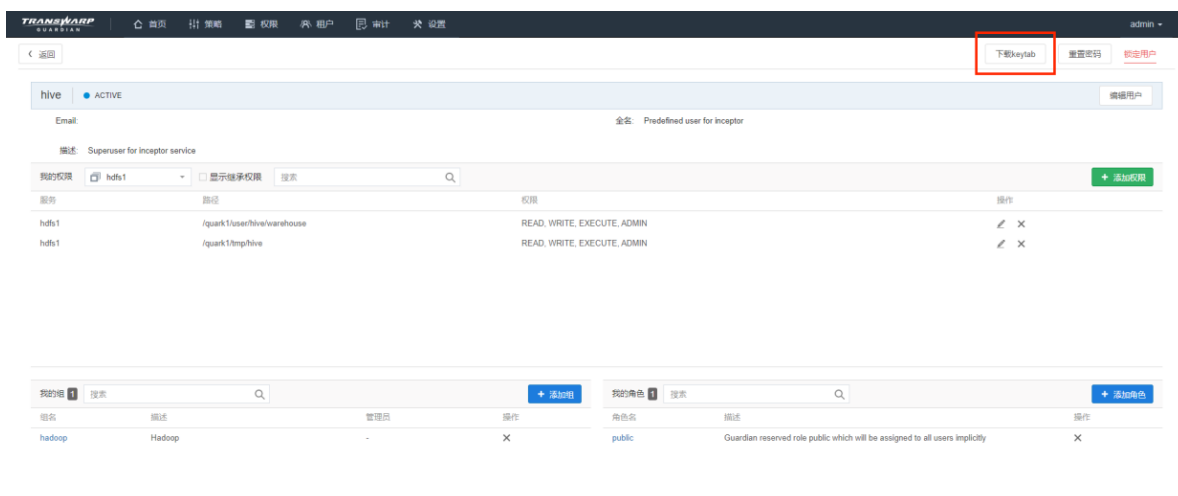
- 步骤一：查看Guardian 详情页中的角色链接



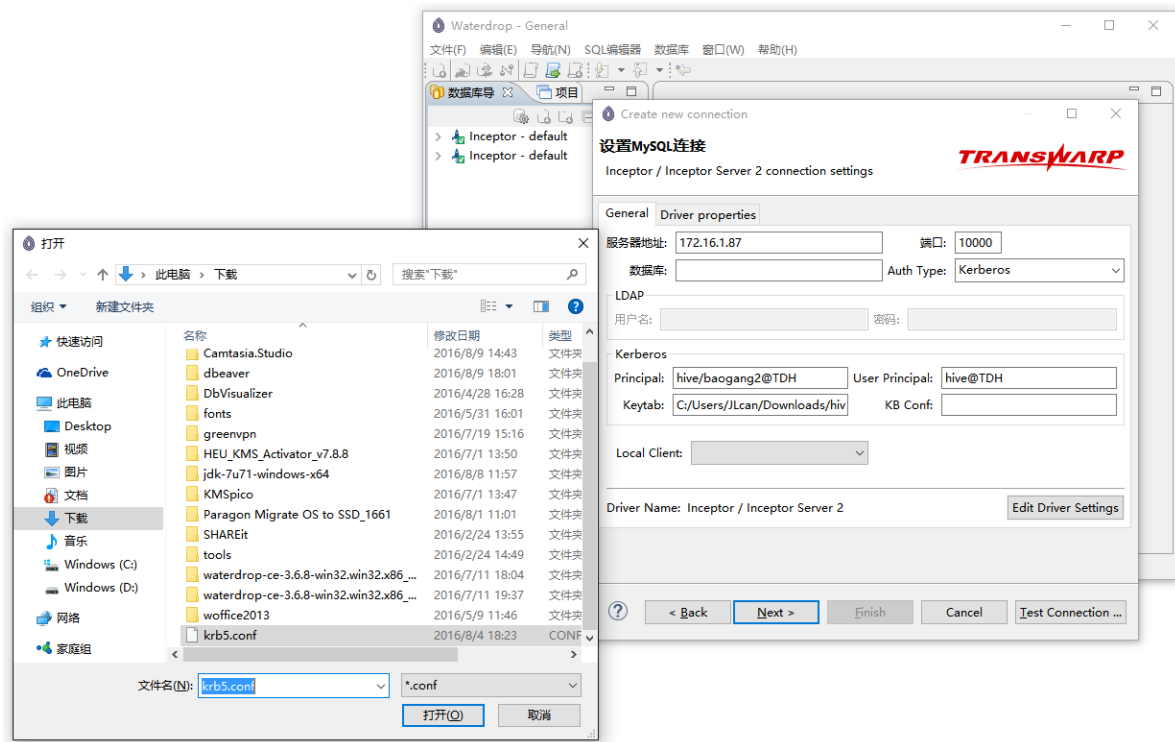
- 步骤二：进入Guardian页面，查看租户列表



- 步骤三：进入对应的租户详情页，点击下载keytab按钮即可获取keytab文件



- 步骤四：找到对应文件的本地位置，双击选中即可。但是需要注意的是放置后如果移动该文件位置则需要重新调整链接配置tab文件。



8.6.7. waterdrop 突然变卡顿是什么原因

【用户问题】

waterdrop 在以前用的时候是不卡顿的，但是这段时间特别卡

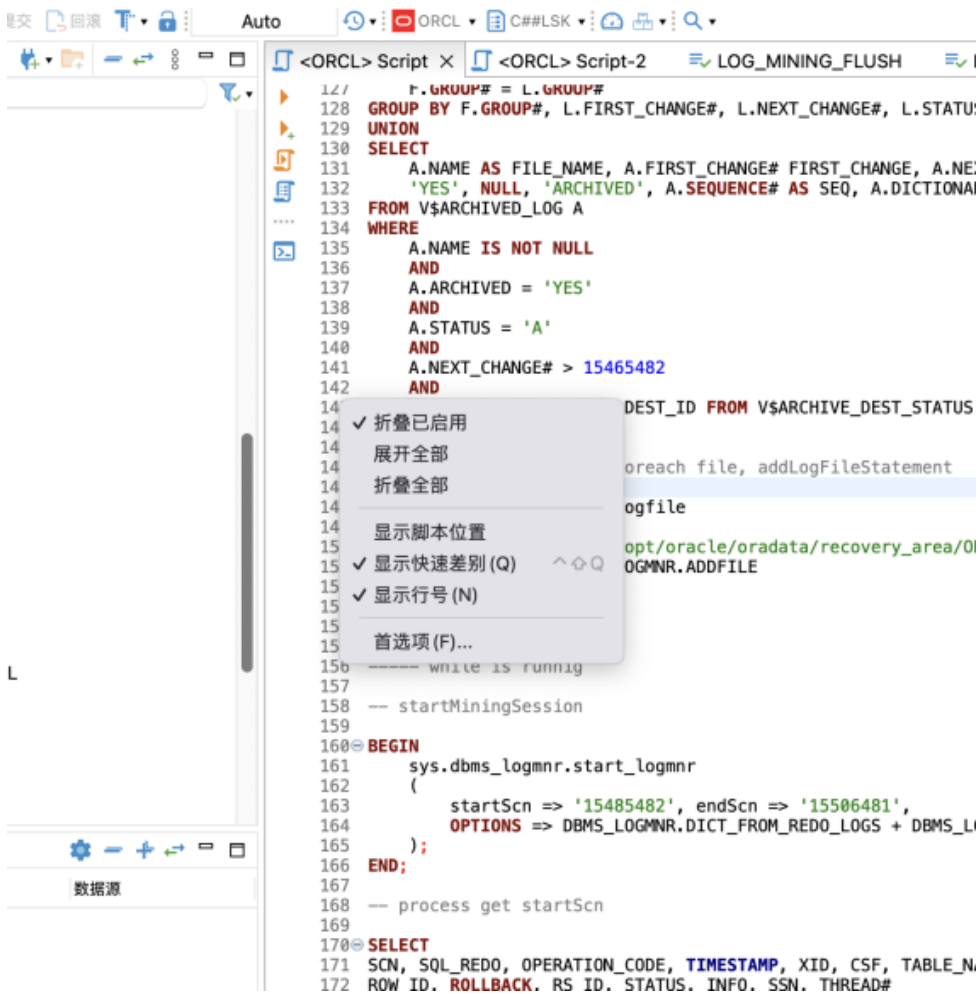
【问题答复】

可能是内存占满了，可以查一下CPU 内存之类的

8.6.8. waterdrop可以设置行号吗?

【问题答复】

可以，参考下图



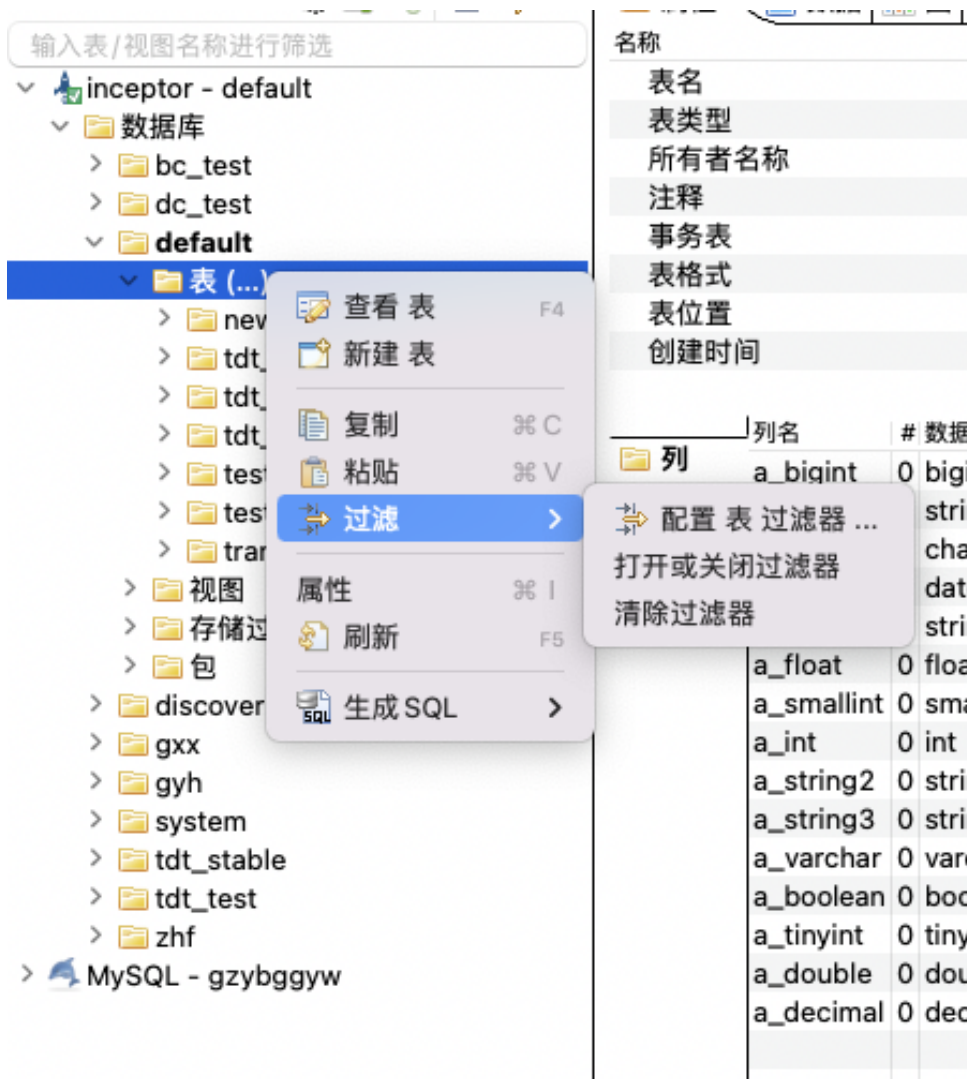
8.6.9. waterdrop-inceptor: filter表如何恢复

【用户问题】

我用waterdrop链接了inceptor，在界面左侧数据表目录中，对一个表右键点击了” filter表 “，还没执行完就点了取消。结果发现左侧目录中这张表隐藏了，sql窗口中可以正常查询这张表。请问要如何恢复这张表的显示？

【问题答复】

参考下图



8.6.10. Waterdrop如何把数据库中表的数据导出到本地?

【用户问题】

使用waterdrop如何把数据库中表的数据导出到本地？导出压缩或为Excel文件

【问题答复】

Waterdrop允许将表中的数据导出至指定文件或者任意连接的表中。方法如下：

1. 在数据库导航栏中选中源表，右键点击，选择导出数据。
2. 选择数据存储格式
3. 设置数据提取方式
4. 设置export属性
5. 选择存放目标数据的地址
6. 确认信息

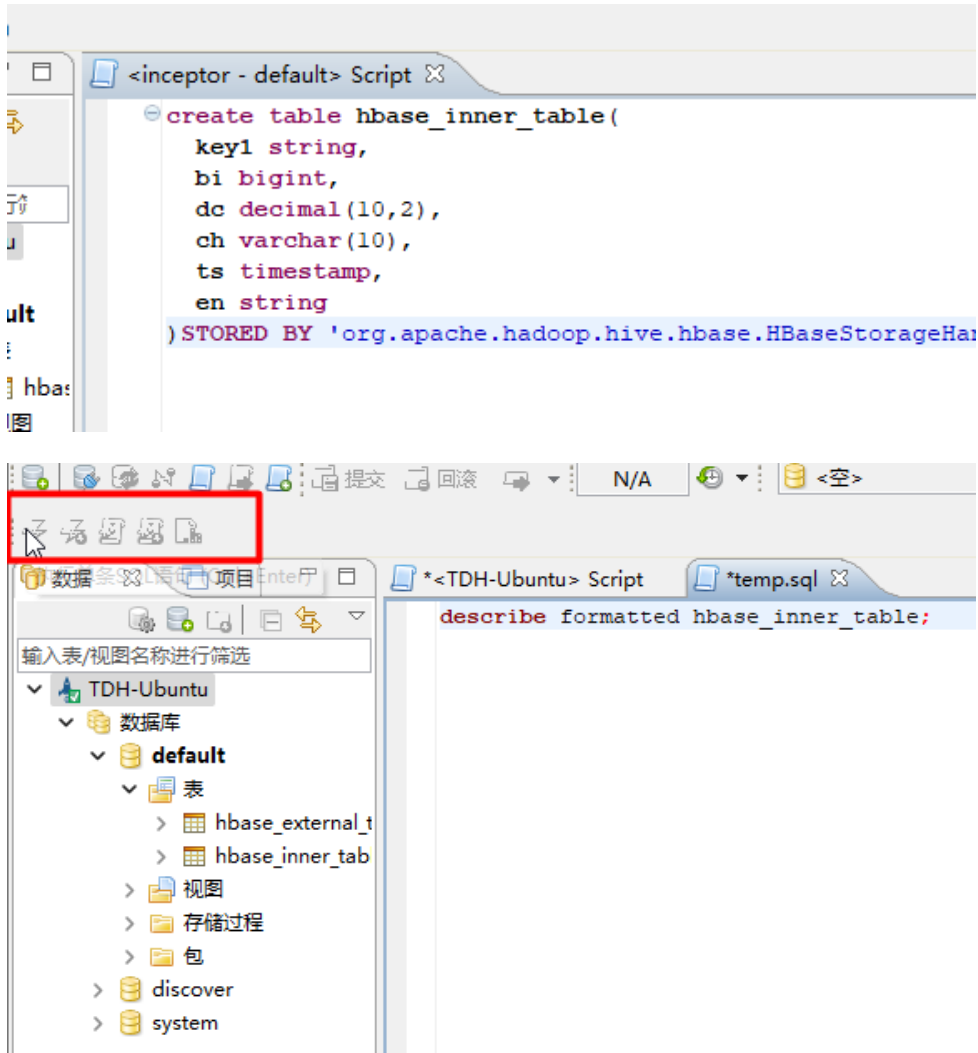
具体详细信息可参考Waterdrop使用手册，第8.2章节

8.6.11. 基于Waterdrop进行数据开发时，SQL Script如何存储以及导入新的SQL脚本?

【用户问题】

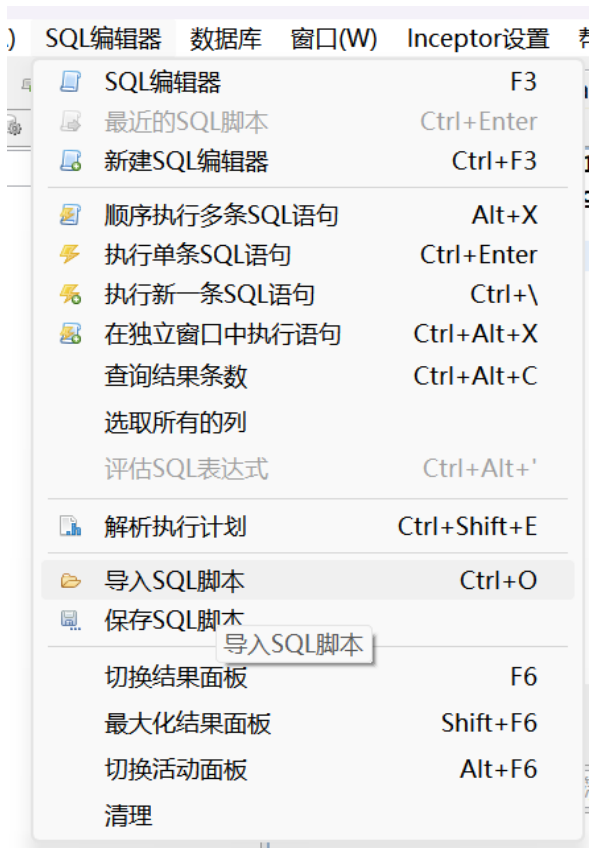
请问我在 Waterdrop 中新建的 SQL script, 在windows哪个文件夹中能找到?

另外我从别的地方拖进来的 SQL script: "temp.sql", 不能直接执行, 请问这应该怎样与数据库建立关联?



【问题答复】

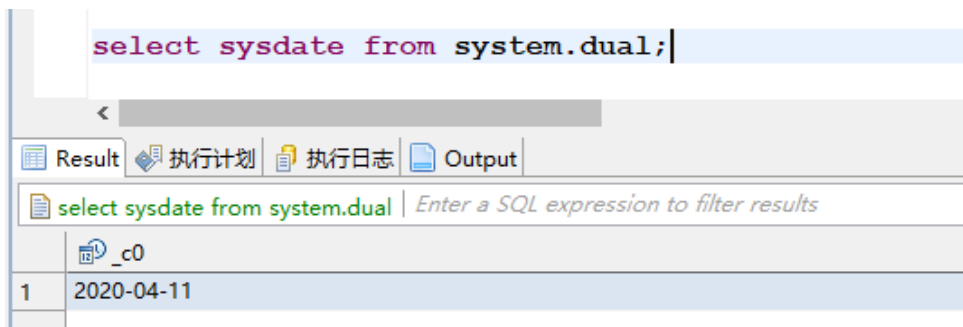
您好, 在保存已编辑完成的SQL script时, 为自定义保存路径, 由您自行选择。连接数据库后, 通过SQL编辑器—导入SQL脚本, 选择您想导入的目标脚本文件。



8.6.12. Waterdrop调用SYSDATE函数返回结果没有时分秒

【用户问题】

在waterdrop中，调用 SYSDATE 函数，返回结果没有时分秒。



【详细说明】

SYSDATE函数说明

8.8.8.22. SYSDATE

函数名	返回类型	描述
SYSDATE	DATE	返回当前系统时间

说明

返回的时间是 'yyyy-MM-dd HH:mm:ss' 格式。注意使用时函数后没有括号。

举例

```
SELECT SYSDATE AS result FROM system.dual LIMIT 1;
```

result
2018-09-10 19:54:23

beeline中调用SYSDATE函数返回结果正确

```
0: jdbc:hive2://172.18.122.113:10000/default> select sysdate from system.dual;
+-----+
|          _c0          |
+-----+
| 2024-01-02 13:27:02 |
+-----+
1 row selected (0.748 seconds)
```

【解决方案】

1. 修改waterdrop设置:

【导航栏 → 窗口 (W) → 首选项 → 一般 → 结果集 → 数据格式】

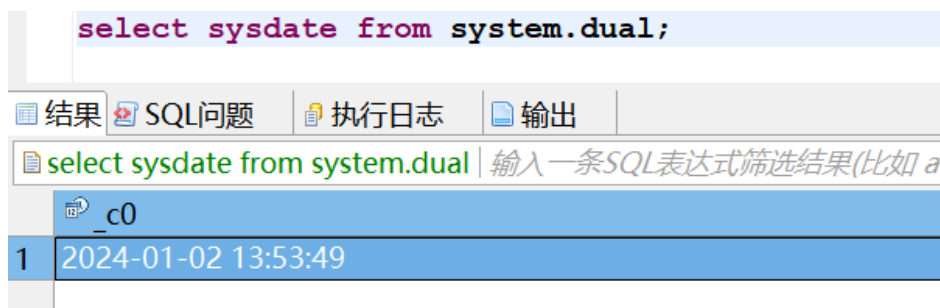
类型 选择 日期

修改 设置 的 值 一栏，将模式改为： yyyy-MM-dd HH:mm:ss

点击 应用 ， 点击 OK 完成



- 再次调用SYSDATE函数，正常显示时分秒。



8.7. 数据库连接相关问题

8.7.1. 如何连接数据库？

相关教程链接：<https://community.transwarp.cn/article/328>

8.7.2. 请问principal的名称如何获取？

【问题描述】

下方框框内的的principle应该在哪里查看？

10.2.3.2.3. Kerberos认证的InceptorServer 2的alias

1. 点击alias列表上方的“+”号来添加一个新的alias。
2. 在这个窗口中您需要：
 - 在Name处填写该alias的名称。这里我们填写hive22。
 - 在Driver处点开下拉条，选中之前我们创建的驱动Hive2。
 - URL处目前显示了我们之前提供的JDBC连接串范例，需要将其改写成实际的JDBC连接串，也就是说要将hostname改成您想要连接的Inceptor server所在节点的IP，并且在URL中提供连接所用的principal、对应keytab和krb5.conf的所在地址，这里以principal为 **hive/CT-1@TDH**，keytab的所在地址为 **E://hive.keytab**；authentication=kerberos，krb5.conf所在地址为 **E://krb5.conf** 为例，提供一个URL示例：

```
jdbc:hive2://172.16.2.65:10000/default;principal=hive/CT-1@TDH;kuser=hive;keytab=E://hive.keytab;authentication=kerberos;krb5conf=E://krb5.conf
```



- hive.keytab从8180的对应用户上下载
- krb5.conf，从/etc下获取

【问题答复】

kuser与principal是一致的；

通常，<principal_name>: Server Principal 中的 User Name 固定为 hive，即本参数的格式为 hive/<Quark 服务所属设备的主机名>@TDH，例如 hive/idc3@TDH。但是前提是用户未自定义过。

具体也可以使用Klist确认一下principle，比如：

```
klist -kt /etc/quark1/conf/quark.keytab
```

8.7.3. 集群外连接不上，报错“could not open connection to jdbc: xxx”怎么排查

【报错信息】

```
[root@alpha-banber-woa-com howdata]# kinit -kt hive.keytab hive@TDH
[root@alpha-banber-woa-com howdata]# klist
Ticket cache: FILE:/tmp/krb5cc_07c7add7-b5bb-4350-b081-b668cfd9ac8
Default principal: hive@TDH

Valid starting Expires Service principal
07/04/2023 19:09:54 07/05/2023 19:09:54 krbtgt/TDH@TDH
renew until 07/11/2023 19:09:54
[root@alpha-banber-woa-com howdata]#
[root@alpha-banber-woa-com howdata]# beeline -u "jdbc:hive2://9.134.77.98:10000/default;principal=hive/9.134.77.98@TDH"
WARNING: log4j.properties is not found, HADOOP_CONF_DIR may be incomplete.
WARNING: Use "yarn jar" to launch YARN applications.
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
scan complete in 12ms
Connecting to jdbc:hive2://9.134.77.98:10000/default;principal=hive/9.134.77.98@TDH
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://9.134.77.98:10000/default;principal=hive/9.134.77.98@TDH: GSS initiate failed (state=,code=0)
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://9.134.77.98:10000/default;principal=hive/9.134.77.98@TDH: GSS initiate failed (state=,code=0)
[root@alpha-banber-woa-com howdata]# beeline -u "jdbc:hive2://9.134.77.98:10000/default;principal=hive/9.134.77.98@TDH"
```

【问题答复】

把beeline命令行记得hive/ip@TDH改为hive/hostname@TDH

8.7.4. 使用beeline命令连接Inceptor时，报错“JAVA_HOME /usr/lib/jvm/java-xxx/bin/java does not exist”是什么原因

【报错信息】

```
[root@tw-node2125 /usr/lib/jvm/jre-1.6.0-openjdk.x86_64/bin]# klist
Ticket cache: FILE:/tmp/krb5cc_780579a0-dd7b-401c-b633-8c89c4720359
Default principal: hive/tw-node2125@TDH

Valid starting Expires Service principal
06/27/2023 12:02:46 06/28/2023 12:02:46 krbtgt/TDH@TDH
renew until 07/04/2023 12:02:46
[root@tw-node2125 /usr/lib/jvm/jre-1.6.0-openjdk.x86_64/bin]# beeline -u "jdbc:hive2://9.134.77.98:10000/default;principal=hive/tw-node2125@TDH"
Unable to determine Hadoop version information.
'hadoop version' returned:
WARNING: log4j.properties is not found. HADOOP_CONF_DIR may be incomplete.
ERROR: JAVA_HOME /usr/lib/jvm/java-1.6.0-openjdk-1.6.0.41.x86_64/jre/bin/java does not exist.

[root@tw-node2125 /usr/lib/jvm/jre-1.6.0-openjdk.x86_64/bin]#
```

【问题答复】

Java环境要求需要是jdk1.8

8.7.5. 使用beeline命令连接Inceptor时，报错“-bash: beeline: command not found”

【报错信息】

```
[root@ce12 ~]# kinit admin
Password for admin@TDH:
[root@ce12 ~]# beeline -u "jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce12@TDH"
-bash: beeline: command not found
```

【问题答复】

在执行beeline连接inceptor之前，需要初始化客户端。

使用命令source /root/TDH-Client/init.sh

8.7.6. 使用beeline命令连接Inceptor时，报错“GSS initiate failed”

【报错信息】

```
[root@ce12 TDH-Client]# beeline -u "jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce12@TDH"
WARNING: log4j.properties is not found. HADOOP_CONF_DIR may be incomplete.
WARNING: Use 'yarn jar' to launch YARN applications.
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
Connecting to jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce12@TDH
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce12@TDH: GSS initiate failed (state=,code=0)
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce12@TDH: GSS initiate failed (state=,code=0)
```

【问题答复】

显示初始化错误，仔细查看对应的安全模式以及链接方式。不同的认证模式连接方式不同，该报错大概率是因为在执行beeline命令前，没有执行kinit <用户名>获取许可证

8.7.7. 使用beeline连接串进入命令行失败，报错显示‘Peer indicated failure: Unsupported mechanism type PLAIN (state=,code=0)’

【问题描述】

链接beeline连接串失败，显示Peer indicated failure: Unsupported mechanism type PLAIN (state=,code=0)

```
[root@ce24 ~]# beeline -u "jdbc:hive2://172.18.130.12:10000/default;principal=hive/ce24@TDH" --maxWidth=15000
WARNING: log4j.properties is not found. HADOOP_CONF_DIR may be incomplete.
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
scan complete in 20ms
Connecting to jdbc:transwarp2://172.18.130.12:10000/default;principal=hive/ce24@TDH
Error: java.sql.SQLException: Could not open connection to jdbc:transwarp2://172.18.130.12:10000/default;principal=hive/ce24@TDH: Peer indicated failure: Unsupported mechanism type PLAIN (state=,code=0)
Error: java.sql.SQLException: Could not open connection to jdbc:transwarp2://172.18.130.12:10000/default;principal=hive/ce24@TDH: Peer indicated failure: Unsupported mechanism type PLAIN (state=,code=0)
```

【问题根因】

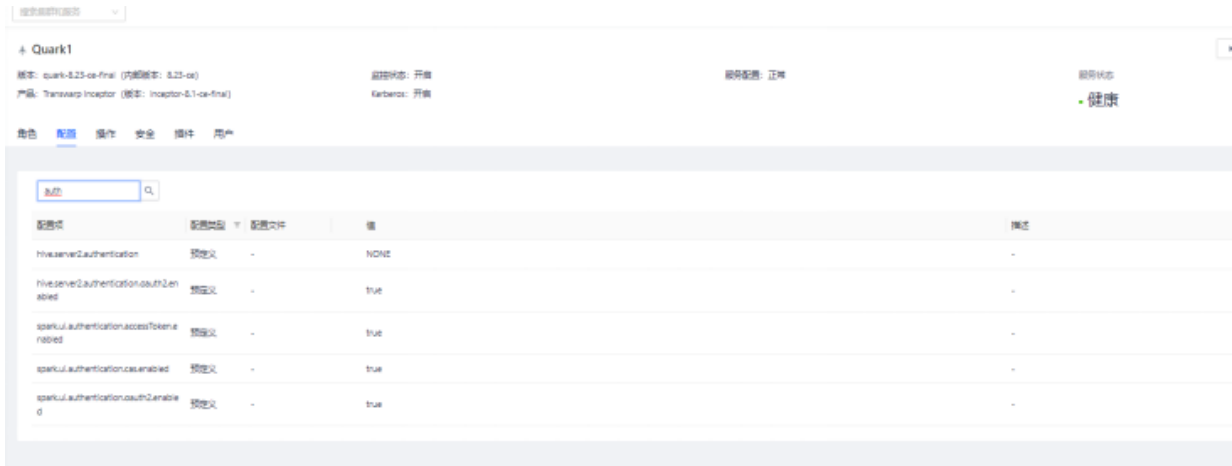
报错是认证方式不对的报错

【相关知识】

开启或未开启安全模式或认证模式不同，beeline连接串的写法也不同，具体详见：
<https://community.transwarp.cn/thread?topicId=276>

【解决步骤】

- 步骤一 配置项输入auth查看认证模式是什么



服务建好后hive.server2.authentication配置项默认为none，未开启安全时可以登录但是如果开启安全的话使用none则无法登录

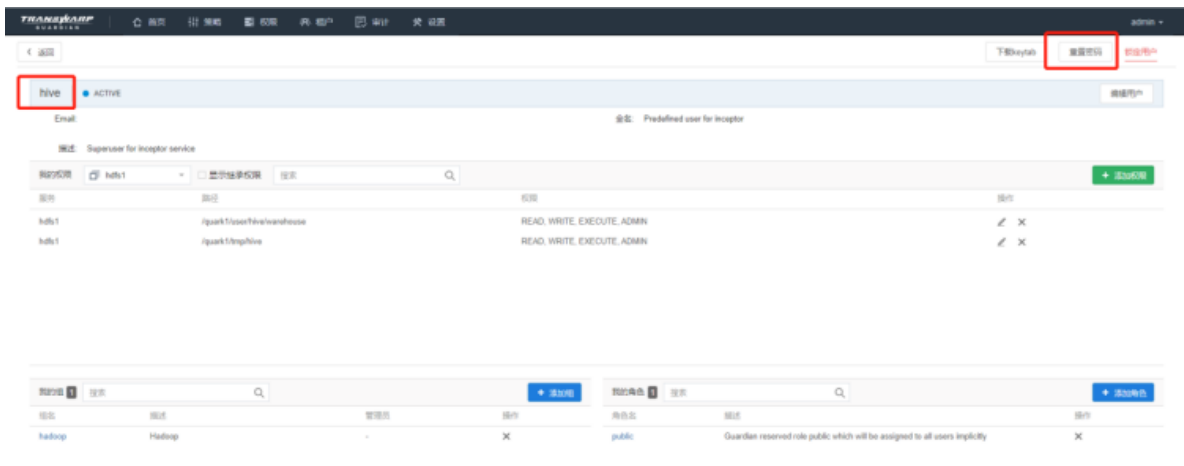
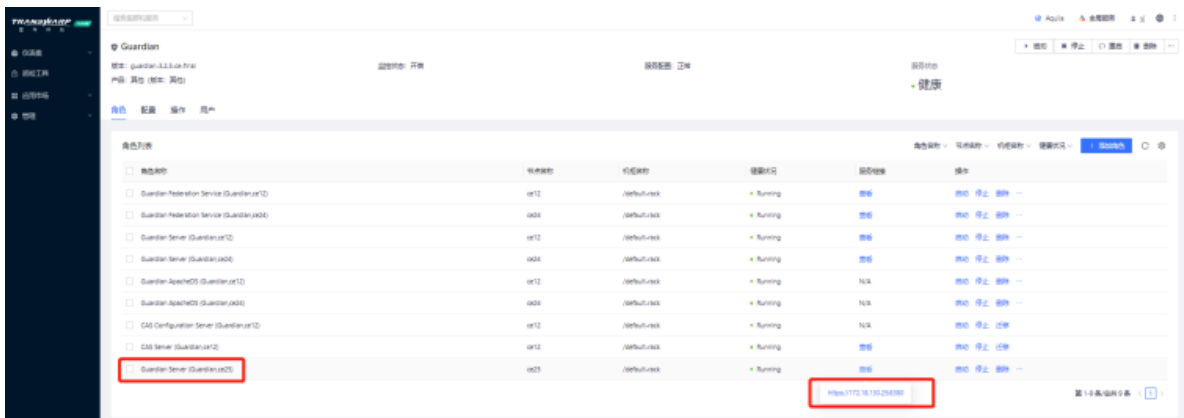
- 步骤二 修改配置项为LDAP，修改后需要重新配置并进行重启

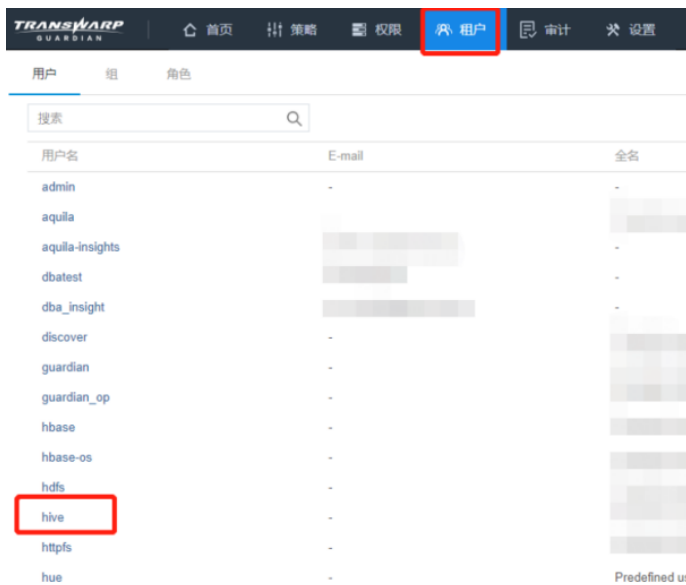


角色 配置 操作 安全 插件 用户

配置项	配置类型	配置文件	值
hive.server2.authentication	预定义	-	LDAP 已修改
hive.server2.authentication.oauth2.enabled	预定义	-	true
spark.ui.authentication.accessToken.enabled	预定义	-	true
spark.ui.authentication.cas.enabled	预定义	-	true
spark.ui.authentication.oauth2.enabled	预定义	-	true

- 步骤三 如果未修改过hive的默认密码，需要到Guardian平台租户页面重置hive密码





- 步骤四 重新使用beeline连接串登录

```
# source /root/TDH-Client/init.sh
# beeline -u "jdbc:hive2://localhost:10000/default" -n <username> -p <password>
```

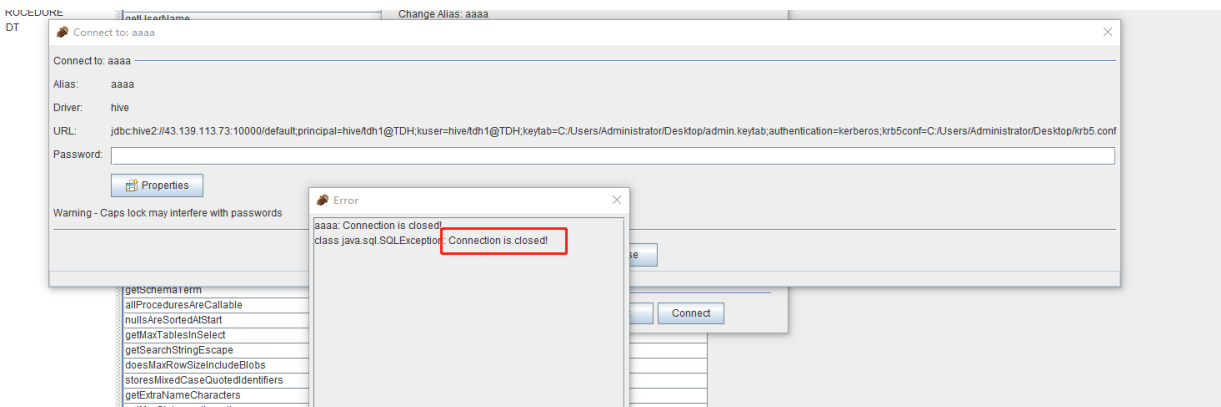
8.7.8. 基于Windows如何使用ODBC配置使用inceptor

【配置教程】

访问链接: <https://community.transwarp.cn/article/965>

8.7.9. Squirrel SQL连接Inceptor时，连接时报错“aaaa:Connection is closed! Class java.sql.SQLException:Connection is closed”

【报错情况】

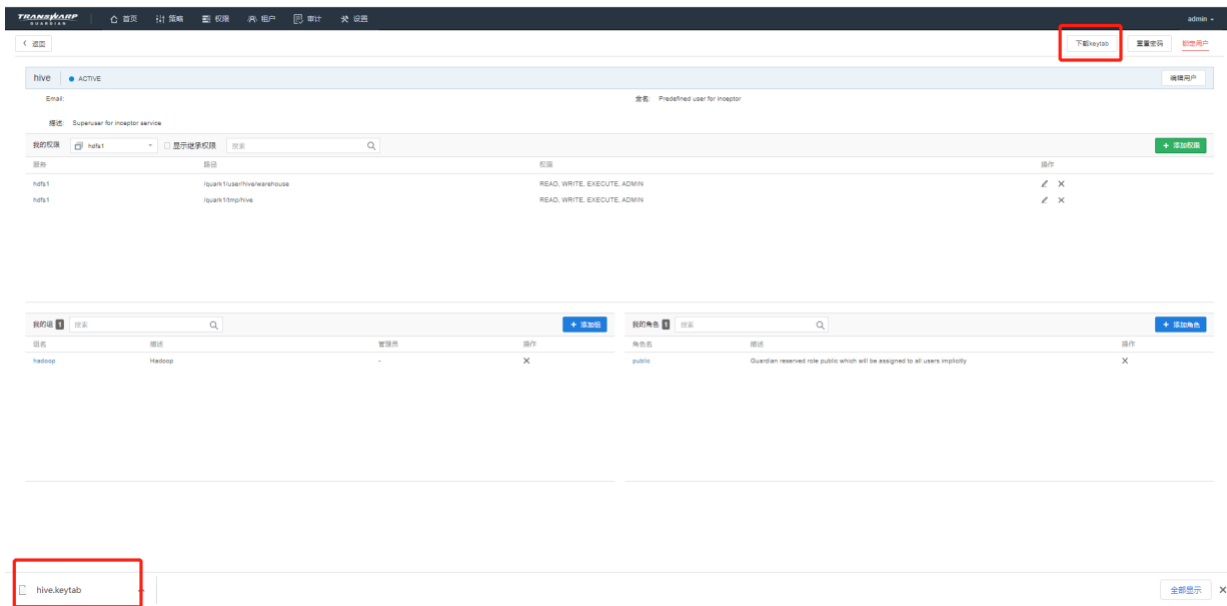
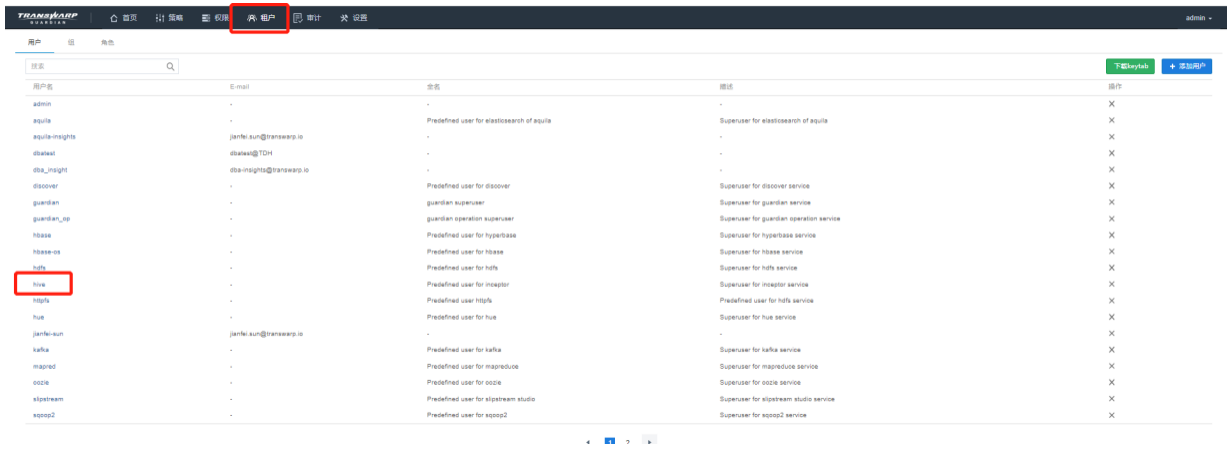


【核心原因】

有可能是因为keytab问题导致

【解决方法】

请下载使用hive.keytab，图片中是admin的，要与principal一致。可以修改下keytab，再连一次。



8.7.10. kerberos登录inceptor数据库， kinit hive的密码多少， 在哪看

【用户问题】

开启了Guardian, inceptor开启的kerberos， 都是默认设置， 在哪查看kerberos的票据的账号密码， 8380上没有可以修改的地方。



【问题答复】

使用下方命令所在节点的hive principal获得一张有效的TGT，以便让Inceptor辨识身份：

```
kinit-kt <hive_keytab_path> hive/<hostname>@TDH
```

这里 <hostname>处提供您所登陆的节点的hostname。<hive_keytab_path>处您提供hive/<hostname>@TDH的keytab文件所在位置。使用示例如下：

```
kinit-kt /etc/sql2/hive.keytab hive/baogang2@TDH
klist
Ticket cache: FILE:/tmp/krb5cc_0
Default principal: hive/baogang2@TDH
Valid starting Expires Service principal 11/21/15 15:27:03 11/22/15 01:27:03
krbtgt/TDH@TDH
renew until 11/22/15 15:27:03
```

这时连接Inceptor时的身份就是hive。连接Inceptor的指令是：

```
beeline -u "jdbc:hive2://<server_ip/hostname>:10000/default;principal=<hive_principal>"。
```

8.7.11. 开启了kerberos后无法采用beeline的方式无法连接数据库，报错这是为什么？

【报错情况】

```
[root@tw-node2125 ~]# beeline -u "jdbc:hive2://9.134.77.98:10000/default;principal=hive/tw-node2125@TDH"
WARNING: log4j.properties is not found. HADOOP_CONF_DIR may be incomplete.
WARNING: Use "yarn jar" to launch YARN applications.
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
scan complete in 25ms
Connecting to jdbc:hive2://9.134.77.98:10000/default;principal=hive/tw-node2125@TDH
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://9.134.77.98:10000/default;principal=hive/tw-node2125@TDH: GSS initiate failed (state=,code=0)
Error: java.sql.SQLException: Could not open connection to jdbc:hive2://9.134.77.98:10000/default;principal=hive/tw-node2125@TDH: GSS initiate failed (state=,code=0)
[root@tw-node2125 ~]#
```

【问题答复】

切换安全模式后，需要重新下载TDH-Client，并进行解压、初始化的操作。此外，如果是开启安全模式，需要基于自身环境判断是Kerberos模式还是LDAP模式，针对不同模式执行不同操作。比如在执行beeline命令前，执行kinit <用户名>获取许可证，不同方式的具体链接教程可参考：

<https://community.transwarp.cn/article/328>

8.7.12. 开启guardian安全且认证模式是LDAP，用kerberos连接串无法连接Quark，为什么？

【报错情况】

11.1.3. 认证模式切换

在Inceptor服务安装结束后，您仍可以在Inceptor服务配置页面上通过开/关Guardian安全和修改hive.server2.authentication参数的值来在四种认证模式间切换。

1. 无认证模式：关闭Guardian安全并且hive.server2.authentication设置为“NONE”。
2. 只开启LDAP认证模式：关闭Guardian安全并且hive.server2.authentication设置为“LDAP”。
3. 只开启Kerberos认证模式：开启Guardian安全并且hive.server2.authentication设置为“KERBEROS”。
4. 开启Kerberos认证和LDAP认证模式：开启Guardian安全并且hive.server2.authentication设置为“LDAP”。

在Transwarp Manager上进入Inceptor服务设置页面，下面说明如何开关Guardian安全和配置hive.server2.authentication参数。

【问题答复】

同时开启Kerberos和LDAP认证模式时，需使用LDAP认证方式，Kerberos会失效。

8.8. 语法&数据库使用相关问题

8.8.1. 使用数据库时报错怎么解决？

有关Inceptor使用相关的语法问题，可以基于报错码查看《[Inceptor错误代码](#)》

8.8.1.1. 示例一. Insert into报错显示 “[Error 10797] Only allow to single insert into Hyperbase/ES/Transaction Orc, other data destination not allowed…”

【报错信息】

```

Connecting to jdbc:hive2://9.134.77.98:10000/default;principal=hive/_HOST@TDH
Connected to: Transwarp Inceptor (version 8.23.0)
Driver: Inceptor JDBC (version 8.31.0)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 8.31.0 by Apache Hive
0: jdbc:hive2://9.134.77.98:10000/default> insert into test_python values (2,2);
Error: COMPILE FAILED: Semantic error: [Error 10797] Only allow to single insert into Hyperbase/ES/Transaction Orc, other data destination not allowed (state=42000,code=10797)
0: jdbc:hive2://9.134.77.98:10000/default>

```

【解决方法】

基于《[Inceptor错误代码](#)》手册，可以看到该报错的原因与修复方法如下：

10.11. 10797 ERROR_10797

编号	错误信息
10797	Only allow to single insert into Hyperbase/ES/Transaction Orc, other data destination not allowed

- 错误原因
只有ORC事务表才支持INSERT VALUES单条插入数据，普通表只能从外部数据源批量导入数据。
- 修复方法
请使用INSERT SELECT或者LOAD DATA等其他方式向非ORC事务表插入数据。

8.8.2. orc表删除列时，报错Replacing columns cannot drop columns for table

【用户问题】

orc表删除列时，报错Replacing columns cannot drop columns for table

【错误原因】

对ORC表使用REPLACE COLUMNS时，替换的列数比原表中定义的列数少。

【修复方法】

禁止对ORC表（包括普通表和事务表）使用 REPLACE COLUMNS。 如果需要删除部分ORC事务表的内容或全部删除可参考下方语法示例：

```
-- 删除非分区表
DELETE FROM table_name WHERE filter_statement
-- 删除单值分区表中一个分区的部分内容
DELETE FROM table_name PARTITION (partition_key = value) WHERE filter_statement
-- 删除范围分区表中一个分区的部分内容
DELETE FROM table_name PARTITION partition_name WHERE filter_statement
```

8.8.3. 在TDH社区版的Inceptor中，使用Insert into插入单条数据失败，请问是什么原因呢？

【用户问题】

未指定表格式创建表格后，使用Insert into插入单条数据失败吗，请问是什么原因呢，应如何解决？如下图所示：

```
0: jdbc:hive2://10.100.100.100:10000/kangll> INSERT INTO stu VALUES("kangll");
Error: COMPILER FAILED: Semantic error: [Error 10797] Only allow to single insert into Hyperbase/ES/Transaction Orc, other data destination not allowed (state=42000,code=10797)
0: jdbc:hive2://10.100.100.100:10000/kangll>
0: jdbc:hive2://10.100.100.100:10000/kangll> INSERT INTO stu SELECT 'John';
1 row affected (10.584 seconds)
0: jdbc:hive2://10.100.100.100:10000/kangll> select * from stu;
+-----+
| name |
+-----+
| John |
+-----+
1 row selected (3.796 seconds)
0: jdbc:hive2://10.100.100.100:10000/kangll>
```

【问题答复】

创建表时，若没有指定存储格式，默认为TextFile，是不支持插入单行或多行的。

这是一个安全防护型的功能禁用：不要做单条插入，不推荐对普通的文件的表做单条的插入。用批处理的方式插入会保证性能和稳定性。

可以使用orc 事务表，支持单条数据的插入。

8.8.4. 存储过程跑一个多小时后就自动停止了，报错 operation cancelled

【解决办法】

可能是operation timeout了或者手动cancel了；如果是timeout导致的可能跟配置有关，需要看这两个配置：

```
hive.server2.idle.operation.timeout
hive.server2.idle.session.timeout
```

其次，beeline ctrl + c 也会导致结束进程

8.8.5. 并发执行插入sql的话会报移动文件错误，这个有解决办法吗？

【报错信息】


```

hive_err.txt
where s618.jjid is null a); EXECUTION FAILED: Task MOVE error HiveException: [Error 20534] copyFiles: error while moving files!!! Cannot move
hdfs://nameservice1/quark1/tmp/hive/hive/781fcd1b-d783-4d09-9abe-4b4cf0dd806/hive_2023-12-12_11-04-21_449_5360756861368689267-35/-ext-10000/000000_0 to
hdfs://nameservice1/quark1/user/hive/warehouse/default.db/hive/t_idx_srs_ippt_idx/000000_0_copy_34; nested exception is java.sql.SQLException: EXECUTION FAILED: Task
MOVE error HiveException: [Error 20534] copyFiles: error while moving files!!! Cannot move hdfs://nameservice1/quark1/tmp/hive/hive/781fcd1b-d783-4d09-9abe-
4b4cf0dd806/hive_2023-12-12_11-04-21_449_5360756861368689267-35/-ext-10000/000000_0 to
hdfs://nameservice1/quark1/user/hive/warehouse/default.db/hive/t_idx_srs_ippt_idx/000000_0_copy_34
at org.springframework.jdbc.support.SQLStateSQLExceptionTranslator.doTranslate(SQLStateSQLExceptionTranslator.java:107)
at org.springframework.jdbc.support.AbstractFallbackSQLExceptionTranslator.translate(AbstractFallbackSQLExceptionTranslator.java:70)
at org.springframework.jdbc.support.AbstractFallbackSQLExceptionTranslator.translate(AbstractFallbackSQLExceptionTranslator.java:79)
at org.springframework.jdbc.support.AbstractFallbackSQLExceptionTranslator.translate(AbstractFallbackSQLExceptionTranslator.java:79)
at org.springframework.jdbc.core.JdbcTemplate.translateException(JdbcTemplate.java:1541)
at org.springframework.jdbc.core.JdbcTemplate.execute(JdbcTemplate.java:393)
at org.springframework.jdbc.core.JdbcTemplate.execute(JdbcTemplate.java:431)
at hygd.rtzs.task.IDXFormulaTask.taskExecute(IDXFormulaTask.java:107)
at hygd.rtzs.task.SystemTask.execute(SystemTask.java:38)
at hygd.rtzs.task.TaskExecute.execute(TaskExecute.java:58)
at hygd.rtzs.mq.handle.ExecuteSpecificJob.lambda$executeParallelTask$0(ExecuteSpecificJob.java:210)
at java.util.concurrent.FutureTask.run(FutureTask.java:266)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:748)
Caused by: java.sql.SQLException: EXECUTION FAILED: Task MOVE error HiveException: [Error 20534] copyFiles: error while moving files!!! Cannot move
hdfs://nameservice1/quark1/tmp/hive/hive/781fcd1b-d783-4d09-9abe-4b4cf0dd806/hive_2023-12-12_11-04-21_449_5360756861368689267-35/-ext-10000/000000_0 to
hdfs://nameservice1/quark1/user/hive/warehouse/default.db/hive/t_idx_srs_ippt_idx/000000_0_copy_34
at org.apache.hive.jdbc.Utils.verifySuccess(Utils.java:250)
at org.apache.hive.jdbc.Utils.verifySuccessWithInfo(Utils.java:234)
at org.apache.hive.jdbc.HiveStatement.execute(HiveStatement.java:408)
at org.springframework.jdbc.core.JdbcTemplate$1ExecuteStatementCallback.doInStatement(JdbcTemplate.java:422)
at org.springframework.jdbc.core.JdbcTemplate.execute(JdbcTemplate.java:381)
... 9 common frames omitted

```

【问题答复】

并发插入用torc会比较好，如果有很高并发插入的情况，在社区版里最好用上kafka，或者用inceptor plus版或者 argodb。

8.8.6. 在跑批中经常出错，取消后重跑又正常，报错显示FetchFailed (BlockManagerIdxxxx...)

【报错信息】

uffle ad	Write Time	Shuffle Write	Errors
			FetchFailed(BlockManagerId(232, executor-t6gj1-1.yb7lw1z.pod.transwarp.local, 458 53, 42791), shuffleId=465399, mapId=11, reduceId=11)
			FetchFailed(BlockManagerId(232, executor-t6gj1-1.yb7lw1z.pod.transwarp.local, 458 53, 42791), shuffleId=465399, mapId=0, reduceId=10)
607.8 KB	4 ms	668.4 KB (0)	ExecutorLostFailure (executor lost)
1167.9 KB	4 ms	665.3 KB (0)	ExecutorLostFailure (executor lost)

【问题答复】

取shuffle结果没取到，可能原因是 executor gc过于严重了，或者是磁盘io过忙

8.8.7. SQL中Select后字段加别名可正常运行，不加则报错

【用户问题】

```
select a1,a2,b1,b2 from aaa
as a left join bbb as b on b.b1=a.a1;
```

报错: error encountered near token 'b' 但是当把SQL改为(字段加上别名)正常:

```
select a.a1,a.a2,b.b1,b.b2
from aaa as a
left join bbb as b on b.b1=a.a1;
```

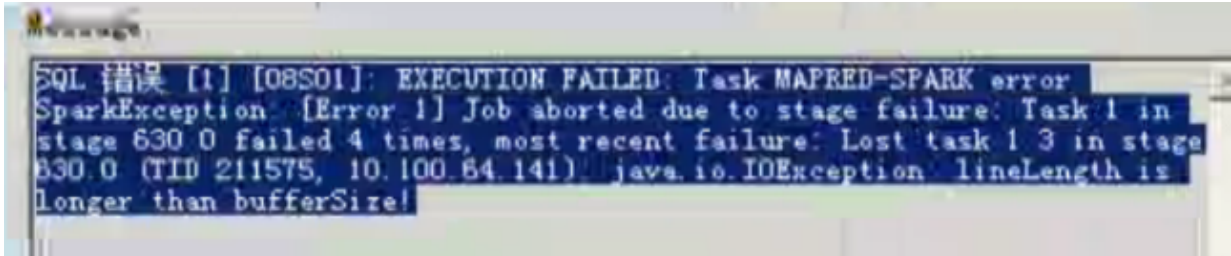
【问题答复】

不可以在select 中不加别名，但在order中加别名，要保持一致就不会出现上述问题。

8.8.8. 查询sql出现错误：lineLength is longer than buffersize，是否有解决方案

【用户问题】

如图，查询时出现一下错误，表格以text方式存储，有string类型的大字段



【问题答复】

这个问题需要配合日志里的callstack看，进一步排错。Callstack需要去quark server或者inceptor-server 所在节点上查看/var/log/SERVICE_ID 比如 /var/log/quark1 或者 /var/log/inceptorX，下面有个quark-server.log 或者inceptor-server2.log

hdfs dfs -rm -skipTrash，报错提示“rm:You are not allowed to do this!”

【用户追问】

这个是什么原因，怎么配置？

【问题答复】

不加skiptrash就可以执行了，skipTrash这个命令被默认禁用了，不可以直接跳过trash，这样会数据丢失哦

【用户追问】

不加是可以执行的，如果基于某些原因必须要加应该如何开启？

【问题答复】

我们是为了防止误删除，故意禁用的，如果一定要加需要先执行：

```
export HADOOP_USER_NAME=hdfs 与
export HADOOP_CLIENT_OPTS="-Dtranswarp.maintenance.only.mode=true"
```

8.8.9. 租户quark的权限都配置了，但是访问的时候还是提示没权限

【报错情况】

The screenshot shows the Transwarp Manager interface for user 'gaea' (ACTIVE). It displays a table of permissions for the 'quark1' role across different databases and tables. Below the table, a terminal window shows an error message: 'Error: COMPILE FAILED: Internal error HiveAccessControlException: [Error 20388] Permission denied: Principal [name=gaea, type=USER] does not have following privileges for operation QUERY [[SELECT] on Object [type=TABLE_OR_VIEW, name=default.yyy]] (state=42000,code=20388)'.

服务	数据库	表	列	权限	操作
quark1	-	-	-	CREATE, SELECT, INSERT, UPDATE, DELETE, ADMIN, ACCESS	✎ ✕
quark1	default	-	-	CREATE, SELECT, ADMIN	✎ ✕
quark1	system	-	-	SELECT	✎ ✕

```

192 rows selected (0.363 seconds)
0: jdbc:transwarp2://tdh-node1:10000/default> select * from yyy;
Error: COMPILE FAILED: Internal error HiveAccessControlException: [Error 20388] Permission denied: Principal [name=gaea, type=USER] does not have following privileges for operation QUERY [[SELECT] on Object [type=TABLE_OR_VIEW, name=default.yyy]] (state=42000,code=20388)
0: jdbc:transwarp2://tdh-node1:10000/default>
0: jdbc:transwarp2://tdh-node1:10000/default>

```

【解决方法】

应该是未配置权限参数，需要登录至 Quark 节点所属设备，将下述配置添加至 `/etc/quark{服务ID}/conf/hive-site.xml` 文件中。

```

<property>
  <name>inceptor.security.column.authorization.enabled</name>
  <value>true</value>
</property>

```

或者也可以登录 Transwarp Manager 平台，在页面左侧，单击仪表盘 > 集群，找到并单击 Quark 服务。单击配置页签，然后单击右侧的添加自定义参数。

The screenshot shows the '添加自定义参数' (Add Custom Parameter) dialog box in the Transwarp Manager. The configuration item is 'inceptor.security.column.authorization.enabled', the value is 'true', and the configuration file is 'hive-site.xml'.

在右侧弹出的面板中，设置配置项为 `inceptor.security.column.authorization.enabled`，取值为 `true`，配置文件为 `hive-site.xml`，然后单击确定。

更具体的如何配置权限流程可参考该demo教程：<https://community.transwarp.cn/article/969>

8.8.10. 用insert overview方法把一个orc表的数据导出生成数据文件，导出时没有分区，重新load进去时如何指定分区？

【用户问题】

我用insert overview方法把一个orc表的数据导出生成数据文件，文件是单值分区，共两个分区，每个分区11个分桶，共22个文件，在导出时没有分区，现在想重新将数据load进去我应该如何指定具体的分区呢？

【问题答复】

您之前那个新表是什么样的 也是跟源表一样的分区么？

【用户答复】

完全一样

【问题追答】

您可以用hdfs dfs -cp -r 这个命令操作；如果说insert overwrite后出来的表格式是txt，跟您的目标表可能格式不一样的话，可以用insert into select的形式，但是需要select 除了分区那一列的其他列，然后后面过滤条件where那里加上分区键。

【用户追问】

这个方法本质其实和get应该差不多吧，都是直接把表的文件完整的拿下来对吗

【问题答复】

嗯嗯，本质上差不多，看您那边主要是想看怎么直接把数据导回新表，那可能最直接的方式就是直接cp回去。

【用户追问】

有示例吗？

【问题答复】

```
CREATE TABLE user_single_date (name STRING, money INT) PARTITIONED BY (reg_date DATE) CLUSTERED BY (money) INTO 3 BUCKETS STORED AS ORC ;

ALTER TABLE user_single_date ADD PARTITION (reg_date = "2021-11-11");
ALTER TABLE user_single_date ADD PARTITION (reg_date = "2021-11-12");
ALTER TABLE user_single_date ADD PARTITION (reg_date = "2021-11-13");

insert into user_single_date partition (reg_date="2021-11-11") select "amber", 100000000 from system.dual;
insert into user_single_date partition (reg_date="2021-11-12") select "lyra", 100000000 from system.dual;
insert into user_single_date partition (reg_date="2021-11-12") select "yijie", 100000000 from system.dual;

CREATE TABLE user_single_date2 (name STRING, money INT) PARTITIONED BY (reg_date DATE) CLUSTERED BY (money) INTO 3 BUCKETS STORED AS ORC;

ALTER TABLE user_single_date2 ADD PARTITION (reg_date = "2021-11-11");
ALTER TABLE user_single_date2 ADD PARTITION (reg_date = "2021-11-12");
ALTER TABLE user_single_date2 ADD PARTITION (reg_date = "2021-11-13");

//LOAD DATA INPATH '/tmp/lyra' INTO TABLE user_single_date2 PARTITION (reg_date="2021-11-11");

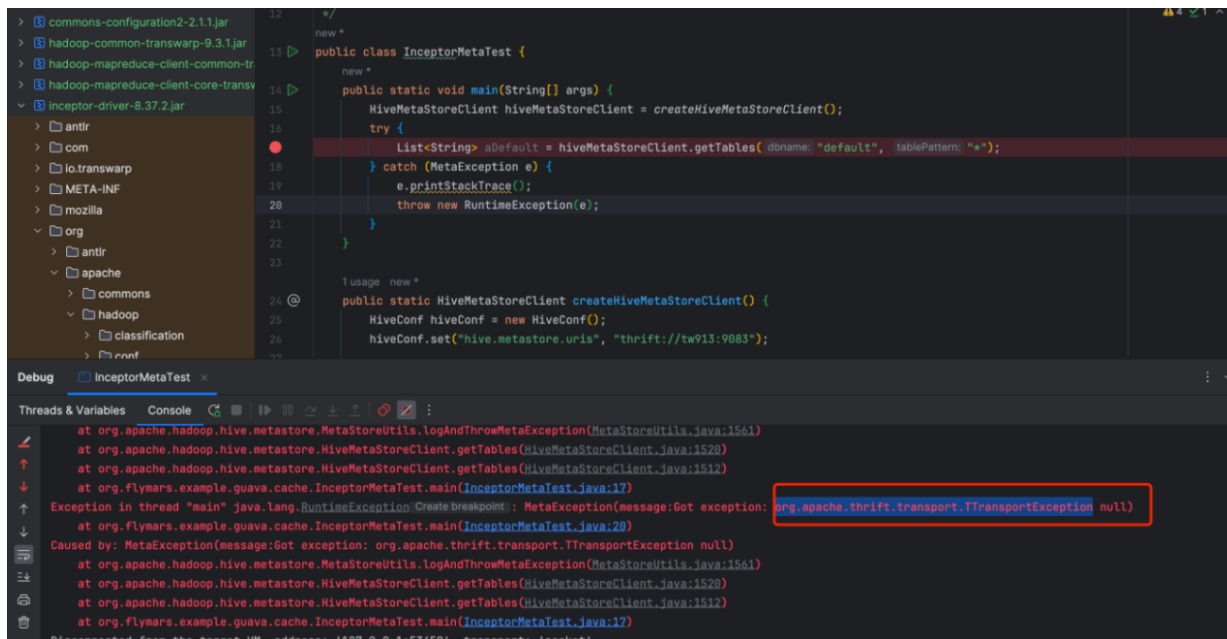
INSERT INTO user_single_date2 PARTITION (reg_date="2021-11-11") SELECT name, money FROM user_single_date WHERE reg_date="2021-11-11";
INSERT INTO user_single_date2 PARTITION (reg_date="2021-11-12") SELECT name, money FROM user_single_date WHERE reg_date="2021-11-12";

//TXT
INSERT OVERWRITE DIRECTORY '/tmp/lyra' ROW FORMAT DELIMITED FIELDS TERMINATED BY '|' select * from user_single_date;

INSERT OVERWRITE DIRECTORY '/tmp/lyra' select * from user_single_date; //txt
```

8.8.11. Inceptor通过使用HiveMetaStoreClient获取表信息的时候出错，应该怎么处理呢

【报错信息】



【问题答复】

这个接口Inceptor原则上是不维护的，如果需要获取表、字段的元数据信息可以查看这篇文章：<https://community.transwarp.cn/article/detail?id=301>

8.8.12. 如何获悉索引信息？

【用户问题】

有没有索引信息的获取方法？

【问题答复】

可以使用 SELECT HOLO_INDEX 查看指定数据库表中的索引信息。语法格式：

```
SELECT HOLO_INDEX ("<database_name>", "<table_name>");
```

参数说明：

<database_name>: 数据库名称;
<table_name>: 表名称;

示例. 查看表 single_fgindex_t 中的索引

```
SELECT HOLO_INDEX ("demodata", "single_fgindex_t");
```

查询结果如下：

column_name	table_name	database_name	index_type
TRANS_ID	single_fgindex_t	demodata	global

8.8.13. 如何统计inceptor中某一表的使用频率？

【问题答复】

目前暂时不支持直接查询。可以使用程序读取日志，统计 select 指定表的次数，该方案较繁琐，可联系售前/售后咨询解决。若您对数据库操作、用户登录等信息有较高的收集监测需求，可以了解一下星环监测与审计软件Audit，实时监测平台中的针对敏感数据的操作事件。

8.8.14. 如何获取一个数据库中存储的所有表元数据信息，包括表名，字段，表数据条数，表大小，表空间等等？

【问题答复】

首先星环没有表空间的概念。

如果需要查看行数可以通过analyze查看，ANALYZE TABLE table_name COMPUTE STATISTICS; desc formatted 就能看到。

如果需要获悉表的元数据信息及其他相关信息，可以连接上数据库后用数据字典来查，比如查看system数据库中的tables_v 和table_stats_v 这两张系统表，需要字段信息则可以查看columns_v，不过最好不要全量查，加个过滤条件比如database_name = xxx。

操作示例：

```
SELECT *  
FROM system.table_stats_v  
WHERE database_name = 'xx'
```

【相关链接】

如需了解星环字段级系统表参考列表可查看：<https://community.transwarp.cn/article/301>

- **table_stats_v**

table_stats_v 系统表存储了 ArgoDB 中所有hdfs表的大小。如果不是hdfs表，则 'table_size' 字段为空。

Table 14. table_stats_v 的字段

字段名称	类型	描述
table_name	varchar2(128,ORACLE)	表名称
database_name	varchar2(128,ORACLE)	表所属数据库名称
table_format	string	表格式
table_location	varchar2(4000,ORACLE)	表的存储位置
table_size	string	hdfs表的大小

- **tables_v**

tables_v 系统表存储了 ArgoDB 中所有表的相关信息。

Table 15. tables_v 的字段

字段名称	类型	描述
table_id	bigint	表编号
table_name	varchar2(128,ORACLE)	表名称
database_name	varchar2(128,ORACLE)	表所属数据库名称
create_time	timestamp	表的创建时间
table_type	varchar2(128,ORACLE)	表类型
owner_name	varchar2(767,ORACLE)	表所属者的名称（一般为创建者）
commentstring	string	表的注释信息
transactional	string	是否为事务表
last_load_time	timestamp	最后加载时间
input_format	varchar2(4000,ORACLE)	表的输入格式
table_format	string	表格式
table_location	varchar2(4000,ORACLE)	表的存储位置
row_permission	string	行级权限
column_permission	string	列级权限
hbase_name	string	hbase 名称
field_delim	string	表中列的分隔符
line_delim	string	表中行的分隔符
collection_delim	string	表中集合列的分隔符

8.8.15. Inceptor可以设置主键自动增长吗？

【问题答复】

分布式下主键自增跟单机数据库不一样，可以用sequence之类的或者窗口函数row number，前者递增不一定连续，性能稍好于后者

8.8.16. inceptor没有distinct on (字段)去重的语句，如何创建视图（取最新的）

【用户问题】

gp有一个distinct on (字段)去重的语句，而inceptor没有，只有distinct 字段。如果要以柜员号为唯一字段创建视图（如下，取最新的），像inceptor应该怎么写？


```
select distinct on (id) *  
from table  
order by id, create_time desc
```

【问题答复】

可以通过窗口函数，在窗口函数内排序，然后算一个排名出来，最后把每个partition 排名第一的选出来

8.8.17. 怎么生成UUID?

【用户问题】

我想select结果，都带一个新的uuid，应该怎么搞？

【问题答复】

有几种方法，性能最好的是用uniq() 函数，但是uniq 函数有一些限制，任务出错重试还可能导致重复数据，也不能用于分桶值，之类的；sequence性能次之，递增不保证连续；row_number 性能最差，递增连续；可以考虑用uniq() 函数

8.8.18. 怎么计算wordcount?

【问题答复】

主要围绕下方4个步骤，更详细的操作步骤请前往社区查看：<https://community.transwarp.cn/article/256>

- a. 创建wordcount.txt文件，并上传HDFS
- b. 创建内表，并导入HDFS数据
- c. 创建结果表
- d. 计算wordcount

8.8.19. 分桶按多少数据量分比较好?

【用户问题】

如果只有100多万条数据，以后增长的也不多，分桶分几个好，100万一个？

【问题答复】

分3个5个都是可以的，100M-200M指的是压缩以后的数据量不是压缩前的。

8.8.20. insert into批量插入数据，会有行数限制吗?

【用户问题】

使用Inceptor通过insert into table values (1,xxx), (2,xxx);

这种方法批量插入数据，一般多少行限制，例如10000行。

是否有更合适批量插入数据的sql命令呢？

【问题答复】

批量插入数据推荐使用 BATCHINSERT 语法实现批量插入功能，此外，您还可以开启 BATCHINSERT RBO 优化，通过直接写入文件的方式获得批量写入性能的极大提升。详情请参考《社区版开发者指南8.2章节》

8.8.21. altertable怎么添加orc表的字段?

【问题答复】

```
alter table post_range_partition_orc2_4 add columns(co11 date);
```

8.8.22. 如何查看执行查询的时候有没有走索引

【问题答复】

quark server日志里可以查看这些关键字，表示走了index scan

```
2022-11-02 18:54:29,543 INFO client.HyperbaseTuples: [InceptorServer2-Handler-Pool: Thread-1672()] - convert globalIndex ws_order_number's indexScan={"filter":"FusionRowFilter
er use next key hint: 0. BytesRanges: [BytesRange offset:0 length:5 start bytes:0 -128 0 17 92 .stop bytes:0 -128 0 17 92 ]","startRow":"\\x00\\x00\\x00\\x11\\x5C","stopRow":
"\\x00\\x00\\x00\\x11","batch":-1,"cacheBlocks":true,"totalColumns":0,"maxResultsSize":-1,"families":{"","cachng":-1,"maxVersions":1,"timeRange":["0","9223372036854775807"]
}
}
2022-11-02 18:54:29,544 INFO client.ScanConstructor: [InceptorServer2-Handler-Pool: Thread-1672()] - Hyperdrive auto select index in local mode, index name is ws_order_numbe
r
2022-11-02 18:54:29,550 INFO task.IteratableFetchTask: [InceptorServer2-Handler-Pool: Thread-1672()] - IteratableFetchTask::IterinitIteratorAndInitLocalEnv took 23 ms
2022-11-02 18:54:29,562 INFO zookeeper.HyperbaseMetadataZKListener: [InceptorServer2-Handler-Pool: Thread-1672()] - start to register zk listener for hyperbase metadata, zk
node : /hyperbase1/hyperbasemetadata/default/web_sales_hyperdrive
2022-11-02 18:54:29,566 INFO zookeeper.ZooKeeper: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e()] - Initiating client connection, connectString=vrsp13
9:2181,vrsp140:2181,vrsp147:2181 sessionTimeout=180000 watcher=org.apache.hadoop.hbase.zookeeper.ReadOnlyZKClient$Lambda$81/619627364065d3c95f
2022-11-02 18:54:29,569 INFO zookeeper.ClientCnxnSocket: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e()] - jute.maxbuffer value is 4194304 Bytes
2022-11-02 18:54:29,571 INFO zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e()] - zookeeper.request.timeout value is 0. feature ena
bled=
2022-11-02 18:54:29,574 INFO zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e-SendThread(vrsp140:2181)()] - Expect server principal:
zookeeper/vrsp140
2022-11-02 18:54:29,574 WARN zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e-SendThread(vrsp140:2181)()] - SASL configuration fail
ed: javax.security.auth.login.LoginException: No JAAS configuration section named 'Client' was found in specified JAAS configuration file: /etc/quark1/conf/jaas.conf'. Will c
ontinue connection to Zookeeper server without SASL authentication, if Zookeeper server allows it.
2022-11-02 18:54:29,574 INFO zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e-SendThread(vrsp140:2181)()] - Opening socket connectio
n to server vrsp140/172.18.123.140:2181
2022-11-02 18:54:29,575 INFO zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e-SendThread(vrsp140:2181)()] - Socket connection establ
ished, initiating session, client: /172.18.123.139:48830, server: vrsp140/172.18.123.140:2181
2022-11-02 18:54:29,579 INFO zookeeper.ClientCnxn: [ReadOnlyZKClient-vrsp139:2181,vrsp140:2181,vrsp147:2181@0x6f54785e-SendThread(vrsp140:2181)()] - Session establishment co
mplete on server vrsp140/172.18.123.140:2181, sessionId = 0x5d00004e477b0204, negotiated timeout = 180000
2022-11-02 18:54:29,580 INFO fulltextindex.EsConstants: [InceptorServer2-Handler-Pool: Thread-1672()] - discovery.zen.ping.unicast.hosts : null
2022-11-02 18:54:29,581 INFO fulltextindex.EsConstants: [InceptorServer2-Handler-Pool: Thread-1672()] - cluster.name : null
2022-11-02 18:54:29,581 INFO client.HyperbaseHTable: [InceptorServer2-Handler-Pool: Thread-1672()] - elasticsearch conf missing!
2022-11-02 18:54:29,583 WARN convert.StargateTableConverterFactory$: [InceptorServer2-Handler-Pool: Thread-1672()] - Can not get StargateStoreTableConverter for router hyper
2drive return HDFS StargateStoreTableConverter as default
2022-11-02 18:54:29,584 INFO serde.HyperdriveSerDe: [InceptorServer2-Handler-Pool: Thread-1672()] - [stargate] [query,dml:serde] [stargate] [performance] [query,dml:serde] I
nitialize total spend time 1263 us, initialize InceptorSerDe spend time 5 us, initialize StoreSerDe spend time 689 us, other spend time 569 us.
2022-11-02 18:54:29,593 INFO client.HyperbaseHTable: [InceptorServer2-Handler-Pool: Thread-1672()] - getScanQuery secondaryIndexScanQuery : index_family: "ws_order_number"
index_scan {
  start_row: "\000\200\000\021\1"
  stop_row: "\000\200\000\021"
  filter {
    name: "org.apache.hadoop.hbase.filter.FusionRowFilter"
    serialized_filter: "\n\022\b\000\020\005\032\005\000\200\000\021\1\1\005\000\200\000\021\1"
  }
  time range {
```

8.8.23. 星环大数据平台存储中文占据几个字节?

【用户问题】

存储到数据库中一个中文字符占据几个字节呢?

【问题答复】

在内存中算三个字节，但是物理存储上不同的存储格式不一样

【用户追问】

oracle数据库编码是gbk的数据，字段类型和长度是varchar2(60)，如果要存到星环数据库，需要扩充长度吗？星环数据库的编码是utf-8。

【问题答复】

要扩充，可以考虑加倍，主要是关系型数据库的设计，和分布式数据库不太一样，分布式数据库的字段精度增加不会特别影响性能

8.8.24. 用日期分区，只能用范围分区吗?

【用户问题】

如果我每天都有几百万的数据，我可以按照天来增加一个分区写入吗？

【问题答复】

也可以使用动态分区。具体可以参考手册中的 8.8 章节，有动态分区的使用介绍和案例：

<https://docs.transwarp.cn/#/documents-support/docs-detail/document/TDH-PLATFORM/9.3/010Inceptor-Developer-Guide?docType=docs&docName=Inceptor%E4%BD%BF%E7%94%A8%E6%89%8B%E5%86%8C>

8.8.25. 有什么办法查询分区表某一分区的大小么?

【问题答复】

是什么表呢? 如果是HDFS表的话可以直接去看hdfs目录大小, 先desc formatted 看到location, 然后
hdfs dfs -du -h -s xxxx

8.8.26. 向分区表导入数据的时候, 必须要添加partition(分区字段=? ?)条件吗?

【用户问题】

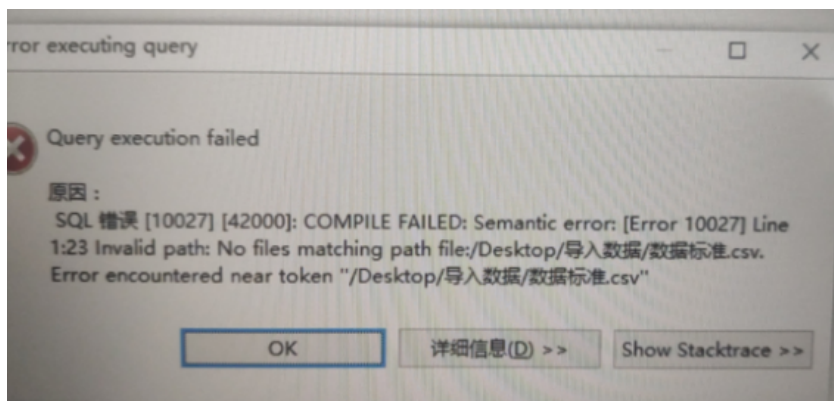
使用load data 的时候, 如果目标表是分区表的话, 需要添加partition(分区字段=? ?)条件吗? 一定要加上分区字段等于某个值吗?

【问题答复】

对的, 需要带上分区的关键字, 无论单值还是范围分区表, 用自己分区表的表现形式就行; 单值是column=value, 范围是partition part_name

8.8.27. 本地load csv 到星环表中报错

【报错信息】



【解决方法】

看报错是文件位置的问题, 需要把文件put到hdfs里, 这里的本地指的是inceptor的容器本地



8.8.28. 如何将 csv 导入 hdfs 中?

【问题答复】

CSV 表的数据来源是 CSV 文件。CSV 文件是纯文本文件，文件中包含数据以及分隔符。和 TEXT 表相似，CSV 表最常见的使用场景是用于建外表，将 CSV 文件中的数据导入 Inceptor，星环科技不建议在任何计算场景中使用 CSV 表。

计算时，应该总是将 CSV 表中的数据用 INSERT ... SELECT 语句插入 ORC 或者 Holodesk 表。

【语法】

创建 CSV 外表

```
CREATE EXTERNAL TABLE <table_name>
(<column_name> <data_type>, <column_name> <data_type>, ...)
STORED AS CSVFILE
[LOCATIONION '<hdfs_path>']-- [1]
[TBLPROPERTIES ( --[2]
['field.delim']='<field_delimiter>',] --[3]
['line.delim']='<newline_char>',] --[4]
['serialization.null.format']='<null_pre>',] --[5]
['quote.delim']='<quote_delimiter>',] -- [6]
['<property_name>']='<property_value>',]...); --[7]
```

参数说明:

- <table_name>: 表名称。
- <column_name>: 列名称。
- <data_type>: 数据类型。
- <hdfs_path>: 表的存储路径，可以是相对路径也可以是绝对路径。
- <field_delimiter>: 字段分隔符，默认值为 “,”。由 <field_delimiter> 分隔的字段会被解析为不同列中的字段。
- <newline_char>: 行分隔符，默认为 “\n”。
- <null_pre>: NULL 值字符，默认为空字段。只支持单个字符，即一个 CHAR。
- <quote_delimiter>: 指定用什么字符作为单个字段的分隔符，默认值为 “”。

- <property_name>: 表属性名称。
- <property_value>: 表属性值。
- 可选项[1], 这里用于手动指定表在 HDFS 中的路径。
- 可选项[2], 支持使用 TBLPROPERTIES 设定表的属性。
- 可选项[3], 支持使用 field.delim 指定字段(列)分隔符。
- 可选项[4], 支持使用 line.delim 指定行分隔符。
- 可选项[5], 支持使用 serialization.null.format 定义 NULL 值字段的存储字符, 默认为空字段。
- 可选项[6], 支持使用 quote.delim 指定每列内部单个字段的分隔符。
- 可选项[7], 持上述 可选项 [3] - [6] 之外, 还支持设定表的其他属性信息。

【示例】

csv文件如下: field1,"field2-part1" "field2-part2, field3-part3", field3 2field1,"2field2-part1 2field2-part2, 2field3-part3", 2field3

建表语句如下:

```
CREATE EXTERNAL TABLE csv_table
(
  col1 STRING,
  col2 STRING,
  col3 STRING
)
STORED AS CSVFILE
LOCATION '/user/alice/csv1'
TBLPROPERTIES(
  'field.delim'=',',
  'quote.delim'='"',
  'line.delim'='\n');
```

8.8.29. 安装完的HDFS用户是什么, 如何切换到hdfs的用户上传文件?

【用户问题】

应该怎么指定, 我需要切换到hdfs的用户, 上传文件

【问题答复】

直接在命令行切换超级用户: export HADOOP_USER_NAME=hdfs

8.8.30. 外部表指定一个文件路径后, 若当前路径下有多个文件, 如何处理?

【用户问题】

星环外部表读取指定目录下的文件时, 是否当前目录下所有文件都读一遍, 然后把格式和表结构一致{分隔符也一致}的文件读取?

【问题答复】

指定目录时, 下面的文件都会被认为是表数据

8.8.31. 加载一个文件, 一直说路径不存在

【用户问题】

是什么情况呢, 确认路径是正确的

```

,code=10027)
8: jdbc:inceptor2://192.168.2.25:10000/default> LOAD DATA LOCAL INPATH '/home/hive/jksj/JKSJ_BZ.DAT' INTO TABLE jksj_bz;
Error: COMPILE FAILED: Semantic error: [Error 10027] Line 1:23 Invalid path: No files matching path file:/home/hive/jksj/JKSJ_BZ.DAT. Error encountered near token ''/home/hive/jksj/JKSJ_BZ.DAT'' (state=42000
,code=10027)

```

【问题答复】

如果用LOCAL INPATH，需要把这个文件翻到Quark server POD里的对应路径下。但是不建议这么使用，建议把文件push到hdfs 然后创建外表这样

8.8.32. SQL游标没有内容，循环进不去

【用户问题】

输入下方p1SQL能运行，但是游标没有内容。循环进不去，会不会是游标后面的查询语句有问题？

```

!set plsqlUseSlash true
declare
  cursor cur is select distinct database_name,table_name from system.tables_v
where database_name = 'default';
begin
  for rs in cur
  loop
    dbms_output.put_line(rs.database_name)
    dbms_output.put_line(rs.table_name)
  end loop;
end;

```

【问题答复】

建议刷新试一下，有关游标使用方法及示例，可查看该文章：<https://community.transwarp.cn/article/330>

8.8.33. 用odbc执行生成函数，sql执行 !set plsqlUseSlash true无法处理

【用户问题】

是否!set只能在命令行中执行，不能当作sql执行？

【问题答复】

!set plsqlUseSlash true不算是sql语句，需要在beeline命令行里执行。如果是odbc的话，是不需要!set plsqlUseSlash true的。

8.8.34. 用oracle方言写函数或存储过程，中间不允许有分号吗？

【用户问题】

beeline -u xxxx -f xxx.sql 使用这种方法编译存储过程，sql文件里是不是不允许有分号；否则就会报错说有<EOF>字符？

【问题答复】

可能是slash的问题，解决方法是每次写存储过程时在存储过程文件的第一行添加 !set plsqlUseSlash true。记得先执行beeline再set。

- 2.在使用记录数据类型变量时，需要在声明部分先定义记录的组成、记录的变量，然后在执行部分引用该记录变量本身或其中的成员。
- 3.只能处理单行数据，数据的类型可以自己定义，也可以基于表来定义，只能在PL/SQL 中运行。

```

!set plsqlUseSlash true
DECLARE
TYPE test_record_type IS RECORD
(test_trans_id string, test_trans_time string ,test_price double)
testrecord1 test_record_type
BEGIN
SELECT trans_id,trans_time,price
INTO testrecord1
FROM transactions
WHERE amount=1100;
dbms_output.put_line(testrecord1.test_trans_id)
END;
/

```

【用户追问】

删除函数怎么写？

【问题答复】

drop plsql function就可以。

8.8.35. TDH sql，在多段 union all时，字符类型会自动补充空格

【用户问题】

比如

```

select 'A' from system.dual
union all
select 'ABCD' from system.dual ;

```

得出来的结果是

```

'A      '
'ABCD'

```

请问这种能否通过参数或者升级版本的方式解决呢？

【问题答复】

需要执行下 SET character.literal.as.string=TRUE;

8.8.36. 有间隔符的数据源，如何划分task是多少？

【用户问题】

请问一下如果是数据源是后缀名是dat，但是有间隔符.这种数据源文件的task是多少，是一个还是基于block划分的？

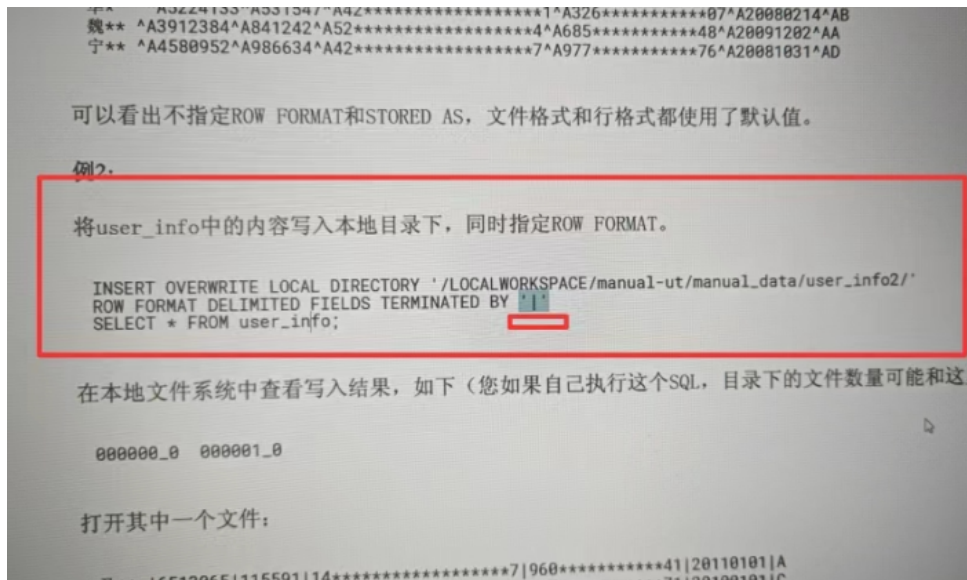
【问题答复】

有明确换行符的 不管是什么后缀，统一都是按照 txt格式处理的，基于block划分task.

8.8.37. TDH-insert overwrite local ...row format delimited_fields terminated by '@'分隔符问题

【用户问题】

在TDH中使用INSERT OVERWRITE LOCAL DIRECTORY 将查询结果导出到本地路径时，设置的分隔符是否可以写成格式为 '@#' 的多字符拼接样式，如果可以为什么实际操作中会数据分隔符号只会读取第一位符号作为分隔符？如果此方式不可以是否有其余方式达到多字符拼接作为分隔符？



【问题答复】

您好，TDH支持使用多字符分隔符做列分隔符或行分隔符，但不支持使用多字符分隔符做 ARRAY，STRUCT，MAP 复杂字段的分隔符。使用 **ROW FORMAT SERDE 'org.apache.hadoop.hive.contrib.serde2.MultiDelimitSerDe'** ... 指定多字符分隔符。

语法格式如下，导出数据语句可类比。

```
CREATE [TEMPORARY] [EXTERNAL] TABLE <table_name> (<column_name> <data_type>, ...)
ROW FORMAT SERDE 'org.apache.hadoop.hive.contrib.serde2.MultiDelimitSerDe'
WITH SERDEPROPERTIES ('input.delimited'='<delimiting_chars>')
[STORED AS TEXTFILE]
[LOCATION '<hdfs_path>']
```

8.8.38. 正在把oracle的程序迁到星环平台，数据排序的顺序是怎么样的？

【用户问题】

我目前有个数据排序的场景，数据中间有null。星环产品null值跟oracle数据库上的null值处理方式一样吗？

【问题答复】

在argodb/inceptor中：

- Oracle和TD方言下，不指定nulls first和nulls last的前提下，null是最小，asc下null在首位，desc下null在末尾；
- DB2方言下，不指定nulls first和nulls last的前提下，null是最大的，asc下null在末尾，desc下null在首位；
- 指定nulls first，desc和asc下，null都在首位；

- 指定nulls last, desc和asc下, null都在末尾;

8.8.39. 表有重复数据发生, 能否使用rowkey来确保唯一?

【用户问题】

请问, 没主键程序控制不严密, 导致表有重复数据发生, 想用rowkey来确保唯一, 这样做会不会对系统性造成影响?

【问题答复】

您好, rowkey的话会做去重排序, 所以数据入库会有个额外操作, 初始导入数据会比普通无脑写入要慢一点点, 但是后面小批量更新之类的操作影响不大。rowkey相关的操作注意事项以及参考可以查看开发者指南8.5章节: <https://docs.transwarp.cn/#/documents-support/docs-detail/document/CE-PLATFORM/23-5/005TDH-CE-Developer-Guide?docType=docs&docName=%E7%A4%BE%E5%8C%BA%E7%89%88%E5%BC%80%E5%8F%91%E8%80%85%E6%8C%87%E5%8D%97>

8.8.40. 函数CURRENT_DATABASE是否可以省略from system.dual?

【用户问题】

星环TDH社区版中, 查询当前数据库是 SELECT CURRENT_DATABASE()可以省略 from system.dual; 而6.2的版本是SELECT CURRENT_DATABASE() from system.dual, 无法省略是吗?

【问题答复】

您好, TDH社区版配有最新版本Inceptor, 对部分函数做出优化, 此处CURRENT_DATABASE函数可以省略from。老版本TDH仍需要填写完整命令。

8.8.41. 请问argodb或inceptor的物化视图日终刷新是怎样的?

【用户问题】

argodb或inceptor数据库的物化视图, 日终结束后刷新逻辑是怎样的呢, 例如执行命令:refresh materialized view v_table; 是需要手动刷新还是会自动刷新呢?

【问题答复】

当view的源表出现更新时, 目前是需要手动刷新物化视图的。该功能正在优化, 后续会支持自动refresh (同步或者异步策略)。

8.8.42. 杀任务, 能通过beeline执行吗?

【用户问题】

页面不方便访问

【问题答复】

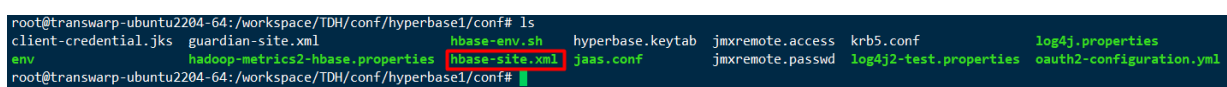
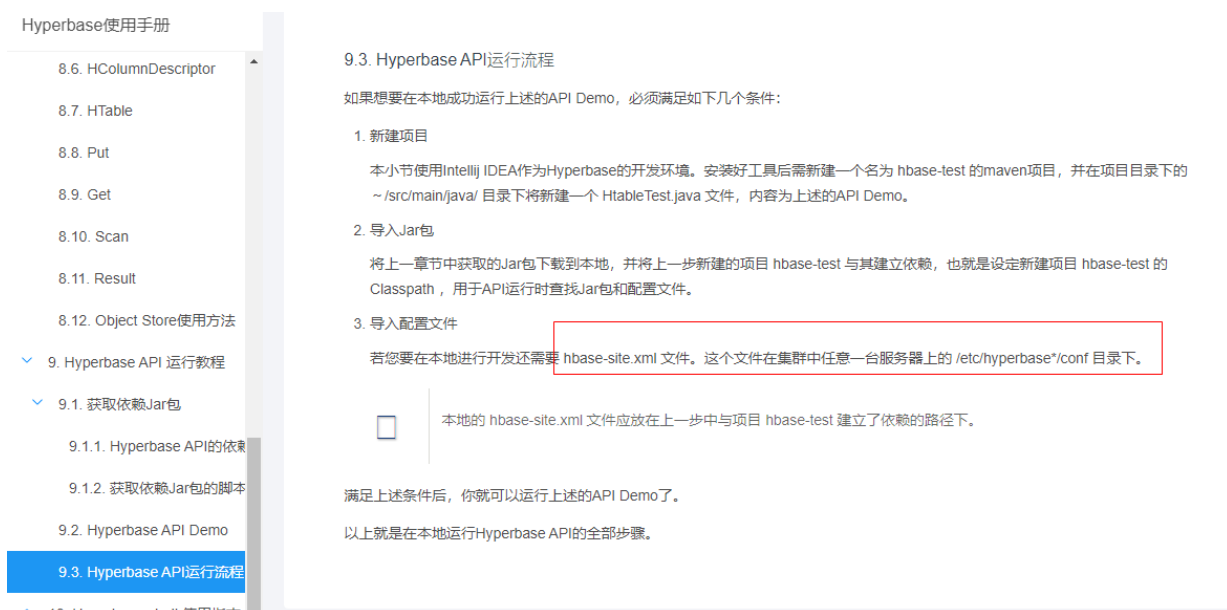
您好, 建议您在DBAService界面对可视化任务进行操作。若操作受限, 可通过beeline来kill sql任务。步骤如下:

1. 查询select * from system.processes_v;来看运行中的sql
2. 通过sessionID和operationID来确定该条sql
3. 执行kill sql operationID

8.8.43. 基于JDBC对Hyperbase数据库进行数据开发时，手册中配置文件如何获取？

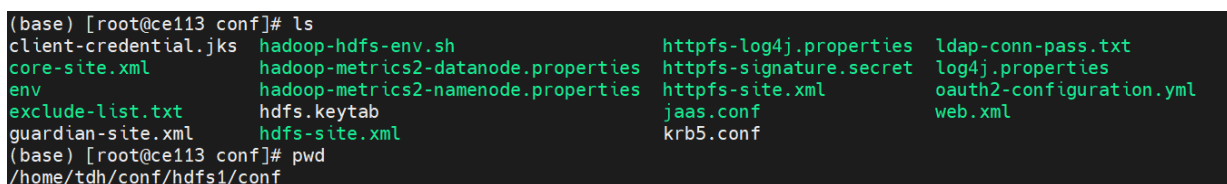
【用户问题】

您好，在使用TDH社区版中，想用 JAVA 代码连接 Hbase，文档的 Demo 中有三个 site.xml 文件，根据文档的路径我只找到了一个 hbase-site.xml，另外两个在哪里找？



【问题答复】

您好，上述配置文件获取路径： /home/tdh/conf/hdfs1/conf



8.8.44. Hyperbase 通过 JAVA API 进行访问，开启安全后如何配置呢？

【用户问题】

在使用社区开发版时，Hyperbase 通过 JAVA API 进行访问，有相关的方法吗？Hyperbase 使用手册中展示未开启 kerberos安全。如开启安全后调用时候会报错没有权限。应如何配置呢？

【问题答复】

可以参考下方链接，该链接介绍了在不同认证模式下，如何使用Java API访问Hyperbase的 Demo，其中展示如何配置krb5.conf等文件。希望对您有所帮助。文章链接：<https://community.transwarp.cn/article/972>

8.8.45. 是否有最新的Hyperbase API Demo

【问题答复】

新的Demo下载链接：<https://transwarp-ce-1253207870.cos.ap-shanghai.myqcloud.com/HyperbaseGovDemo.zip>

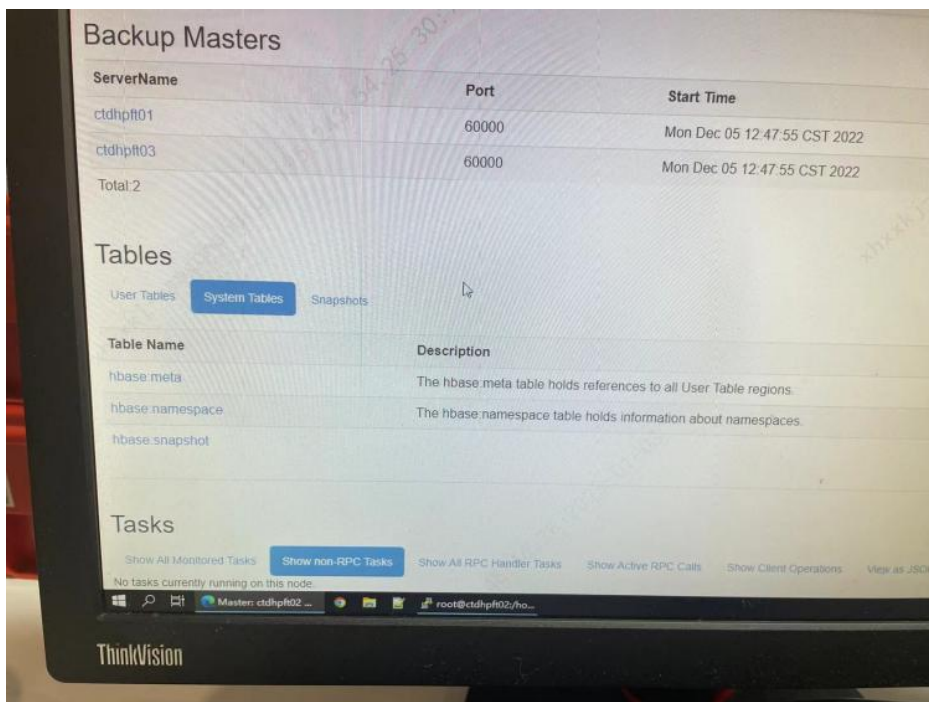
新的Demo添加了配置文件和log4j.properties方便参考，适用于hbase 1.3 和2.2 两个版本，对应我司6.0以上的版本，包括6.0, 6.2, 7.0, 8.0, 8.1, 9.0, 9.1等；

8.8.46. hbase:acl表不存在

【用户问题】

hbase:acl表不存在，该怎么解决啊，是应该这么操作吗？

1. hbase zkcli
2. ls /
3. rmr hyperbase1
4. 重启hyperbase



【问题答复】

您好，您可以选择删除zookeeper里面的hbase:acl对应的ZNODE（如

/hyperbase1/tables/hbase:acl)，然后切换HMaster即可重建。但是可能会存在权限丢失的风险，如果有权限丢失的话在hbase:acl重建后可以重新赋权。

【用户追问】

请问删除ZNODE的命令是这个吗 `delete /hyperbase1/table/hbase:acl`

【问题答复】

是的，您删除/hyperbase1/table/hbase:acl这个znode就可以

【用户追问】

请问切换HMaster，该怎么操作，是重启hyperbase吗？

【问题答复】

您停掉当前状态为active的HMaster就会自动切换。

8.9. 小文件以及合并操作的相关问题

8.9.1. 小文件背景介绍、不同阶段解决方案、不同表格式合并方式等相关链接总览

友情链接

- [为什么会出小文件？小文件出现背景、原因及危害](#)
- [针对不同阶段下的小文件过多问题如何处理，最佳治理手段（星环Compact机制）](#)
- [Text/ORC非事务表合并最佳方式](#)
- [TORC表格式对应的小文件合并方法](#)
- [Inceptor Plus及ArgoDB Holodesk表格式对应的小文件合并方法（Compact Service）](#)
- [TORC Compact 常见故障诊断方法](#)
- [TDH及ArgoDB归档分区功能介绍及使用方法](#)
- [小文件相关问题的处理思路](#)
- [Inceptor compact 性能优化方法](#)
- [Automerge及优化后的AutomergeV2用法及注意事项](#)

8.9.2. 当出现小文件相关的问题以及应该怎么处理

【问题背景】

很多客户在生产环境都会遇到小文件问题，小文件可能来自于上游系统，可能来自于书写不当的sql，也可能是错误数据导致join生成大量小文件。用户侧观察到的现象就是，性能下降，任务报错，甚至executor lost。

【问题分析】

小文件问题存在众多隐患：

1. 导致map 任务非常多，从而性能下降；
2. 小文件过多，可能会影响到后续reduce任务数量暴增，从而导致出现task 超过10w的情况，这时候引擎会出于自我保护会报错；
3. 小文件如果非常小，在进行automerge后，还可能导致executor端，每个task需要读太多小文件，从而会在一个task中新 过多DFSClient，DFSInputStream，Path，JobConf对象等，导致 executor端

内存和gc 压力很大，甚至executor lost。

【排查思路】

一般来说，一个任务是由几个步骤组成的，而小文件的产生也来自任务的各个流程和步骤：

上游 => 本地文件系统 => HDFS => Map => Reduce => FileSink

所以解决小文件问题就是从上面的步骤中入手，并且解决小文件的隐患越早越好，跟过滤下推一样，能尽早做的过滤条件一定是尽早过滤。能在本地文件系统解决的问题，没必要放到HDFS上解决，因为HDFS本就不适合存储大量小文件，小文件过多会导致namenode元数据特别大，占用太多内存，严重影响HDFS的性能。

【解决方法】

1. 上游系统改进
2. linux文件系统合并
3. hdfs 合并：`hdfs dfs -appendToFile a.txt b.txt c.txt hdfs://${namenode}/data/all.txt`
4. concatenate 命令，自动合并orc小文件
5. maptask阶段合并：automerger
 - [Automerger用法及注意事项](#)
 - [改进partition automerger](#)
6. reduce 阶段合并：`reduce tasks` 设置，`mapred.reduce.tasks`，如果没有sql本来没有reduce，可以加上：`distribute by rand()`；
7. filesink合并：
 - [小文件合并改进：](#)
 - [小文件合并设计实现版本](#)

【操作实例】

某用户上游应用产生了57000 个小文件，每个文件一行数据。创建text 外表后，执行任务报了task 超过10w。第一反应先走automerger，开启automerger后，数据本地合并后大概合并了出了不到10 个task，然后就把executor 跑lost了。

问题是每个任务需要读的文件太多了，一个task中连续new DFSClient, DFSInputStream, Path, JobConf 等对象，导致gc压力太大，最后executor lost。

针对这个用户问题最终的解决方案是：

1. 尽量在本地文件系统中合并
2. 本地文件系统如果不能合并，在hdfs 上使用append file 合并
3. 如果不能在local filesystem或者hdfs合并，考虑创建中间表，通过reduce task + distribute by 来降低文件数量
4. 或者使用星环最新的功能Galactus非事务表小文件合并功能：[Text/ORC非事务表合并最佳方式](#)

8.9.3. TORC Compact 常见故障诊断

【现象描述】

正常情况下，系统里orc transaction表大部分都会存在一个base目录，如果发现orc transaction表里有很多delta目录 (>10) 却没有自动compact，则说明该系统已经发生问题了，此时可以按照以下步骤来进

行初步诊断。

【诊断思路】

1. 首先在inceptor里用select语句检查该表能否被正常的select，如果select出错，则问题定位为orc表本身的问题。如果能select出来则说明compact系统出现问题
2. compact系统出现问题时，首先查看compact所在服务的进程的jstack，检查里面是否有以下3种thread:

- a. Initiator thread: 负责检查每个orc transaction表是否符合compact条件
- b. Worker thread: 负责向yarn提交mapreduce任务对表进行compact
- c. Cleaner thread: 负责清除冗余的版本，例如上面的例子中生成base_0246405后delta_0246404_0246404与delta_0246405_0246405就可以被删除掉

如果没有initiator thread，系统不能对表进行自动的compact，但通过alter table还能手动触发compact，如果没有worker thread，系统将无法进行compact，这两种情况都会导致orc transaction表的相关信息在metastore里面积压，大幅影响系统性能甚至会导致系统挂掉。cleaner thread相对比较不重要，但是没有cleaner，hdfs上就会有太多冗余版本，故我们应该保证metastore进程中这三种thread都是处于正常状态，如果发现metastore中少了某个thread，应该重启metastore。

以metastore的jstack为例，这些thread在jstack的表现如下：

```
"Thread-8" daemon prio=10 tid=0x00007f6c4802d000 nid=0x7815 waiting on condition [0x00007f6c43cfb000]
  java.lang.Thread.State: TIMED_WAITING (sleeping)
  at java.lang.Thread.sleep(Native Method)
  at org.apache.hadoop.hive.ql.txn.compactor.Cleaner.run(Cleaner.java:142)

"xufei-ThinkPad-T440p-18" daemon prio=10 tid=0x00007f6c48029800 nid=0x7814 waiting on condition [0x00007f6c43dfc000]
  java.lang.Thread.State: TIMED_WAITING (sleeping)
  at java.lang.Thread.sleep(Native Method)
  at org.apache.hadoop.hive.ql.txn.compactor.Worker.run(Worker.java:95)

"xufei-ThinkPad-T440p-17" daemon prio=10 tid=0x00007f6c48027800 nid=0x7813 waiting on condition [0x00007f6c43ef000]
  java.lang.Thread.State: TIMED_WAITING (sleeping)
  at java.lang.Thread.sleep(Native Method)
  at org.apache.hadoop.hive.ql.txn.compactor.Worker.run(Worker.java:95)

"Thread-5" daemon prio=10 tid=0x00007f6c48026800 nid=0x7812 waiting on condition [0x00007f6c43ffe000]
  java.lang.Thread.State: TIMED_WAITING (sleeping)
  at java.lang.Thread.sleep(Native Method)
  at org.apache.hadoop.hive.ql.txn.compactor.Initiator.run(Initiator.java:149)
```

3. 如果compact所在进程中这三种thread都处于正常状态，对于基于mapreduce的compact，还要检查yarn上compact任务是否正常，之前碰到的情况是yarn分配给inceptor的资源太多，导致compact任务一直无法完成，碰到这种情况需要调整yarn的资源分配来确保有足够的资源进行compact。
4. 对于带分区（partition）的ORC事务表，可以使用以下命令手动触发compaction:

单值分区（single-key partition）

```
alter table single_key_partition_table partition col compact 'major'; //对partition spec
col 进行compact;
alter table single_key_partition_table partition(col='value') compact 'major'; //
对partition spec col=value 进行compact;
```

多值分区表（multi-key partition）

```
alter table multi_key_partition_table partition(col1='value1', col2='value2') compact
'major';
alter table multi_key_partition_table partition col compact 'major'; //对partition spec
col 进行compact;
```


单值范围分区 (single-key and range partition)

```
alter table range_partition_table partition part_key compact 'major'; -- 其中的单引号' ' 是需要的。
```

5. 手动合并失败，查看metastore中的日志报错如下：

```
the object store...
15-23 16:34:13,030 INFO metastore.HiveMetaStore: (HiveMetaStore.java:logInfo(745)) [pool-3-thread-200()] - 200: Metastore shutdown complete.
15-23 16:34:13,039 INFO HiveMetaStore.audit: (HiveMetaStore.java:logAuditEvent(235)) [pool-3-thread-200()] - ugi=hive/whtplocapp0@TUM ip=/10.25.12.131 cmd=Metas
rdown complete.
15-23 16:34:17,209 INFO compactor.Worker: (Worker.java:run(103)) [whtplocapp01.localdomain-68()] - Start check duplicate compact for ci_id = 9343 tbl_obj.t_custom: null
15-23 16:34:17,292 INFO compactor.Worker: (Worker.java:run(119)) [whtplocapp01.localdomain-68()] - End check duplicate compact, ci_id = 9343, needSkip = false
15-23 16:34:17,293 WARN compactor.Worker: (Worker.java:run(122)) [whtplocapp01.localdomain-68()] - tbl_obj.t_custom is in black list, which means compaction for this tab
le has failed 3 times in a row, and compaction is disable. Use alter table tableName enable compact to re-enable compact.
15-23 16:34:17,300 WARN compactor.Worker: (Worker.java:run(130)) [whtplocapp01.localdomain-68()] - tbl_obj.t_custom is in blacklist or another compaction for the same desti
n is on the flight, so skip this one.
15-23 16:34:23,271 INFO metastore.HiveMetaStore: (HiveMetaStore.java:logInfo(745)) [pool-3-thread-200()] - 200: get all databases
15-23 16:34:23,271 INFO HiveMetaStore.audit: (HiveMetaStore.java:logAuditEvent(235)) [pool-3-thread-200()] - ugi=hive/whtplocapp0@TUM ip=/10.25.12.131 cmd=Metas
```

【核心原因】

因为表已经进入了黑名单，进入黑名单的表是不能合并的。

【解决方法】

进入黑名单是因为之前自动合并失败很多次，需要手动执行下述命令从黑名单中释放

```
alter table table_name_xxxxxx enable compact; -- 'table_name_xxxxxx' 替换为表名;
```

同时，为了保证上述语句运行成功，防止delta数量过多的极端情况下合并失败，需要适当调整参数来设置一次性合并的delta版本数量上限：

```
SET hive.compactor.max.num.delta=50; -- 默认值为500
```

8.9.4. Compaction Blacklist机制

【问题描述】

如果一个表或者分区在多次尝试compact并且失败，compaction 服务会认为后续再对这个表或分区compact同样会失败。为了避免浪费资源，compaction服务会将这个表或分区加入compaction blacklist。

目前失败次数由参数orc.compact.blacklist.threshold控制，默认值是3。

表或分区一旦加入黑名单，无论自动或手动触发compaction，都不会执行compact操作。

目前还没有从compaction blacklist中自动移除的机制。

【Compaction中常用的操作方法】

1. 查看黑名单

可以在beeline通过以下命令查看：

```
show compact blacklist
```

2. 从黑名单中移除

如果将某个表或分区从compaction blacklist中移除，可以通过：

在beeline，从blacklist中移除某个表或分区，执行

```
alter table table_name enable compact; // 表
alter table table_name partition (pt='xxx') enable compact; // 单值分区
alter table table_name partition range_name enable compact; // 范围分区
```

也可以在mysql中，直接删除COMPACTION_BLACKLIST中记录

3. 手动加入黑名单

可以手动将某个表或分区加入黑名单，不让其做compaction。

```
alter table table_name disable compact; // 表
alter table table_name partition (pt='xxx') disable compact; // 单值分区
alter table table_name partition range_name disable compact; // 范围分区
```

8.9.5. Insert into产生大量小文件后采用Compact机制不生效，无法提交对应的任务到执行引擎中

【问题描述】

用户在使用jdbc 从另一张表读数据（60w数据量）写入orc事务表中的时候，hdfs上出现了许多delta开头的目录每个目录有5个分桶小文件，小文件过多导致namenode内存不够崩了。

采用Compact机制针对此类情况进行优化后发现没有触发合并，且服务健康。

【核心原因】

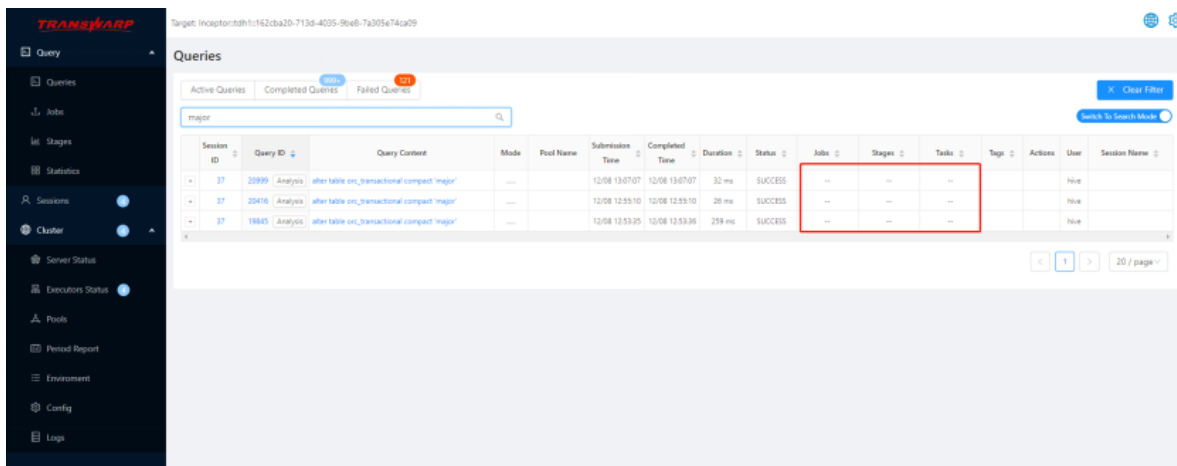
排查了黑名单后发现曾经触发过合并但是并未提交对应的任务到执行引擎中，后排查了metastore日志发现是因为Yarn资源不足导致的。

【排查过程】

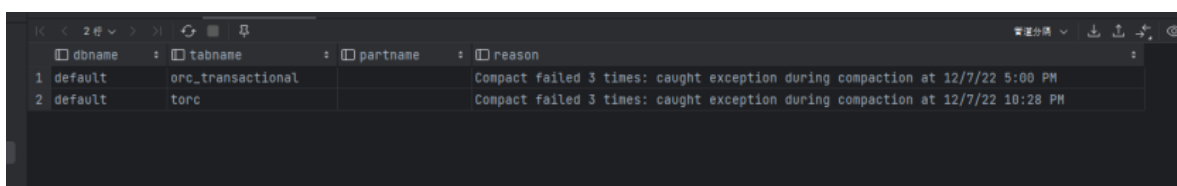
某一社区版用户每执行一次insert into，就会生成一个delta文件

1. 采用compact机制进行合并，metastore这三个线程都正常，但是手动触发compactor并没生成任务跑，当没有base目录时，delta 大于10个文件也没有触发合并；然后设置了metastore 但是并没走yarn 的mr任务；

```
bash-4.2# sudo -u hive ./jstack -l 242 |grep compactor
at org.apache.hadoop.hive.ql.txn.compactor.Cleaner.run(Cleaner.java:191)
at org.apache.hadoop.hive.ql.txn.compactor.Worker.run(Worker.java:282)
at org.apache.hadoop.hive.ql.txn.compactor.Worker.run(Worker.java:282)
at org.apache.hadoop.hive.ql.txn.compactor.Initiator.run(Initiator.java:239)
bash-4.2#
```



2. 排查黑名单，发现表已经进入了黑名单，需要手动从黑名单中释放；

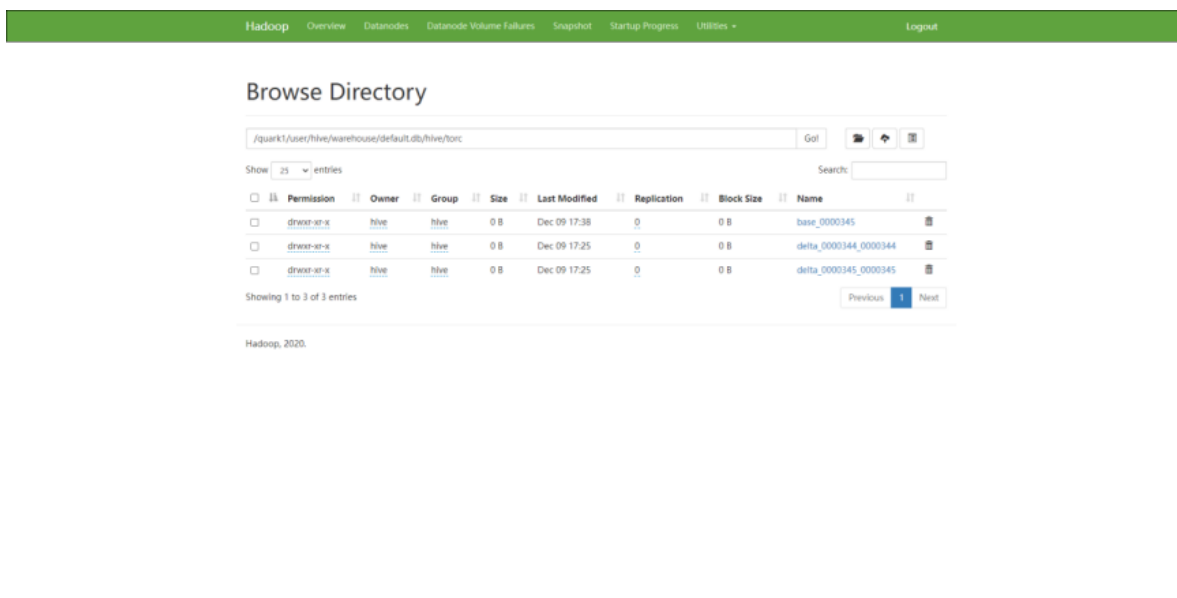


3. 把表从黑名单拿出来后，手动执行依然没反应；

4. 用户需要查看metastore日志进行排查；

`/var/log/quark1/quark-metastore.log`

5. 排查后发现是因为yarn资源不足，改完资源就可以了



```

2022-12-09 17:30:25,960 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=source:172.16.8.211 get_primary_keys : db=default tbl=orc
2022-12-09 17:30:25,965 INFO metastore.InceptorMetaStore: [pool-2-thread-84(1)] - 84: source:172.16.8.211 get_all_databases
2022-12-09 17:30:25,967 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=source:172.16.8.211 get_all_databases
2022-12-09 17:30:25,969 INFO metastore.InceptorMetaStore: [pool-2-thread-84(1)] - 84: source:172.16.8.211 get_packages: db="pat"
2022-12-09 17:30:25,970 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=source:172.16.8.211 get_packages: db="pat"
2022-12-09 17:30:25,973 INFO metastore.InceptorMetaStore: [pool-2-thread-84(1)] - 84: source:172.16.8.211 get table : db=system tbl=dual
2022-12-09 17:30:25,973 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=source:172.16.8.211 get_table : db=system tbl=dual
2022-12-09 17:30:25,980 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=source:172.16.8.211 get_databases: system
2022-12-09 17:30:26,169 INFO metastore.InceptorMetaStore: [pool-2-thread-84(1)] - 84: Shutting down the object store...
2022-12-09 17:30:26,169 INFO metastore.InceptorMetaStore: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=Shutting down the object store...
2022-12-09 17:30:26,170 INFO InceptorMetaStore.audit: [pool-2-thread-84(1)] - ugl-hive ip=172.16.8.211 cmd=metastore shutdown complete.
2022-12-09 17:30:26,244 INFO metastore.ObjectStore: [tbl-40(1)] - ObjectStore, initialize called
2022-12-09 17:30:26,248 INFO metastore.ObjectStore: [tbl-40(1)] - Using direct FS, underlying DB is MySQL
2022-12-09 17:30:26,248 INFO metastore.ObjectStore: [tbl-40(1)] - Initialized ObjectStore
2022-12-09 17:30:26,248 INFO compactor.Worker: [tbl-40(1)] - Start check auto-compact-disabled set for ci_id = 24 default torc null
2022-12-09 17:30:26,250 INFO compactor.Worker: [tbl-40(1)] - Start check duplicate compact for ci_id = 24 default torc null
2022-12-09 17:30:26,251 INFO compactor.Worker: [tbl-40(1)] - End check duplicate compact, ci_id = 24, newComp = false
2022-12-09 17:30:26,258 INFO compactor.Worker: [tbl-40(1)] - Starting MAJOR compaction for default_torc, rp_id = 24
2022-12-09 17:30:26,308 INFO compactor.DrCompactor: [tbl-40(1)] - ValidCompactionList: highWatermark=345, minOpenTime=62233720369054772607, exceptions=[], ctc=[344-305, 345-306]
2022-12-09 17:30:26,337 INFO com.TemHelder: [tbl-40(1)] - 'locks' locked by 'tbl'
2022-12-09 17:30:26,344 INFO com.TemHelder: [tbl-40(1)] - 'locks' unlocked by 'tbl'
2022-12-09 17:30:26,344 INFO com.TemHelder: [tbl-40(1)] - Speed size wait lock lock = 0, handle lock request = 7, cond = 0
2022-12-09 17:30:26,351 INFO compactor.Compactor: [tbl-40(1)] - OK, transactional table directory info: table=default_torc, base=null; deltas=[344, 344], [345, 345]; obsolete=[], sortOrder=[344-344, 345-345]
2022-12-09 17:30:26,351 INFO compactor.Compactor: [tbl-40(1)] - Submitting MAJOR compaction job 'tbl-compactor.default.torc' to default queue. (current delta dirs count=2, obsolete delta dirs count=0, TemIdRange[344,345]
45) Delta dirs are [[344, 344], [345, 345]]
2022-12-09 17:30:26,357 INFO mapreduce.JobResourceCloader: [tbl-40(1)] - Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2022-12-09 17:30:26,406 INFO mapreduce.JobResourceCloader: [tbl-40(1)] - Disabling Erasure Coding for path: /yarn/user/hive/staging/job_1670578603942_0001
2022-12-09 17:30:26,742 INFO mapreduce.JobSubmitter: [tbl-40(1)] - number of splits:3
2022-12-09 17:30:26,744 INFO mapreduce.JobSubmitter: [tbl-40(1)] - Submitting tokens for job: job_1670578603942_0001
2022-12-09 17:30:26,784 INFO mapreduce.JobSubmitter: [tbl-40(1)] - Executing with tokens: []
2022-12-09 17:30:26,729 INFO impl.YarnClientImpl: [tbl-40(1)] - Submitted application application_1670578603942_0001
2022-12-09 17:30:26,774 INFO mapreduce.Job: [tbl-40(1)] - The url to track the job: http://tbl10668/proxy/application_1670578603942_0001/
2022-12-09 17:30:26,779 INFO mapreduce.Job: [tbl-40(1)] - Running job: job_1670578603942_0001
2022-12-09 17:30:26,668 INFO compactor.Worker: [tbl-39(1)] - tbl-39 sleep time prolong to 137838 ms
2022-12-09 17:30:41,418 INFO mapreduce.Job: [tbl-40(1)] - Job job_1670578603942_0001 running in uber mode : false
2022-12-09 17:30:56,532 INFO mapreduce.Job: [tbl-40(1)] - map 0% reduce 0%
2022-12-09 17:30:56,532 INFO mapreduce.Job: [tbl-40(1)] - map 100% reduce 0%
2022-12-09 17:30:57,546 INFO mapreduce.Job: [tbl-40(1)] - Job job_1670578603942_0001 completed successfully
2022-12-09 17:30:57,563 INFO mapreduce.Job: [tbl-40(1)] - Coasters: 32
File System Coasters

```

```

2022-12-09 17:26:14,030 INFO mapreduce.JobSubmitter: [tbl-39(1)] - Submitting tokens for job: job_1670564572967_0003
2022-12-09 17:26:14,031 INFO mapreduce.JobSubmitter: [tbl-39(1)] - Executing with tokens: []
2022-12-09 17:26:14,223 INFO conf.Configuration: [tbl-39(1)] - resource.types.xml not found
2022-12-09 17:26:14,224 INFO resource.ResourceUtils: [tbl-39(1)] - Unable to find 'resource-types.xml'.
2022-12-09 17:26:14,251 INFO mapreduce.JobSubmitter: [tbl-39(1)] - Cleaning up the staging area /yarn/user/hive/staging/job_1670564572967_0003
2022-12-09 17:26:14,220 ERROR compactor.Worker: [tbl-39(1)] - Caught exception while trying to compact default store. Marking clean to avoid repeated failures. java.lang.Exception: Inceptor error: org.apache.hadoop.yarn.exceptions.InvalidResourceRequestException: Invalid resource request, requested resource type=memory:mb < 0 or greater than maximum allowed allocation. Requested resource=memory:1536, vcores:1>, maximum allowed allocation=memory:1024, vcores:2>, please note that maximum allowed allocation is calculated by scheduler based on maximum resource of registered NodeManagers, which might be less than configured maximum allocation=memory:96304, vcores:96>
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.throwInvalidResourceException(SchedulerUtils.java:420)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.checkResourceRequestAgainstAvailableResource(SchedulerUtils.java:402)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.validateResourceRequest(SchedulerUtils.java:311)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:259)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:223)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.validateAndCreateResourceRequest(RMAppManager.java:559)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.createAndPopulateNewRMApp(RMAppManager.java:407)
at org.apache.hadoop.yarn.server.resourcemanager.ClientRMService.submitApplication(RMAppManager.java:342)
at org.apache.hadoop.yarn.server.resourcemanager.ClientRMService.submitApplication(ClientRMService.java:678)
at org.apache.hadoop.yarn.impl.pb.service.ApplicationClientProtocolPBServiceImpl.submitApplication(ApplicationClientProtocolPBServiceImpl.java:277)
at org.apache.hadoop.yarn.proto.ApplicationClientProtocol$ApplicationClientProtocolService$2.callBlockingMethod(ApplicationClientProtocol.java:563)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtobufRpcInvoker.call(ProtobufRpcEngine.java:527)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1036)
at org.apache.hadoop.ipc.Server$RPC$Call.run(Server.java:1015)
at org.apache.hadoop.ipc.Server$RPC$Call.run(Server.java:943)
at java.security.AccessController.doPrivileged(Native Method)
at java.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:2145)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2953)
WR error: org.apache.hadoop.yarn.exceptions.InvalidResourceRequestException: Invalid resource request, requested resource type=memory:mb < 0 or greater than maximum allowed allocation. Requested resource=memory:1536, vcores:1>, maximum allowed allocation=memory:1024, vcores:2>, please note that maximum allowed allocation is calculated by scheduler based on maximum resource of registered NodeManagers, which might be less than configured maximum allocation=memory:96304, vcores:96>
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.throwInvalidResourceException(SchedulerUtils.java:420)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.checkResourceRequestAgainstAvailableResource(SchedulerUtils.java:402)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.validateResourceRequest(SchedulerUtils.java:311)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:259)
at org.apache.hadoop.yarn.server.resourcemanager.scheduler.SchedulerUtils.normalizeAndValidateRequest(SchedulerUtils.java:223)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.validateAndCreateResourceRequest(RMAppManager.java:559)
at org.apache.hadoop.yarn.server.resourcemanager.RMAppManager.createAndPopulateNewRMApp(RMAppManager.java:407)
at org.apache.hadoop.yarn.server.resourcemanager.ClientRMService.submitApplication(RMAppManager.java:342)
at org.apache.hadoop.yarn.server.resourcemanager.ClientRMService.submitApplication(ClientRMService.java:678)
at org.apache.hadoop.yarn.impl.pb.service.ApplicationClientProtocolPBServiceImpl.submitApplication(ApplicationClientProtocolPBServiceImpl.java:277)
at org.apache.hadoop.yarn.proto.ApplicationClientProtocol$ApplicationClientProtocolService$2.callBlockingMethod(ApplicationClientProtocol.java:563)

```

8.10. 自定义函数UDF、UDAF、UDTF相关问题

8.10.1. 如何使用外部jar包中自定义函数UDF、UDAF、UDTF

- 聊一聊UDF/UDTF/UDAF是什么，开发要点及如何使用？：

<https://community.transwarp.cn/article/310>

- UDF部署操作指南（含视频教程）：

<https://community.transwarp.cn/article/929>

8.10.2. 如何创建持久化函数？

【操作步骤】

1. 查看已存在jar包

```
LIST JAR;
```

2. 添加jar包

```
ADD JAR[S] <local_or_hdfs_path>;  
//Local_path是Inceptor server节点的路径。保证hive用户对jar所在的目录有读权限。
```

3. 创建永久UDF

```
CREATE PERMANENT FUNCTION [<db_name>.<function_name>] AS <class_name>;
```

如果Inceptor不在local mode，那么资源的地址也必须是非本地URI，比如HDFS地址。

4. 验证永久UDF

```
SELECT [<db_name>.<function_name>()] FROM SYSTEM.DUAL;
```

更多信息可查看：

UDF部署操作指南（含视频教程）：<https://community.transwarp.cn/article/929>

8.10.3. UDF自定义函数一次支持多少数据使用

【用户问题】

请问UDF自定义函数一次支持多少数据使用？使用UDF创建一个SM4加密解密函数，一次性解密5w以上的数据就会出现问题。

【问题答复】

Inceptor对UDF自定义函数处理的峰值性能可能受到多种因素的影响，其中包括内存大小、网络传输效率、函数实现效率等。针对SM4加密解密需注意CPU占用过高，并且由于加解密通常不适合批量操作，所以一般建议采用分批次处理的方式，避免一次性处理过多数据造成的问题。例如：数据分批处理；对大型表进行分区，然后分区分表调用UDF；

8.10.4. UDF函数调用时报错（一）

【用户问题】

写了个调用http接口的udf函数，注册为临时函数，调用时报错：com.esotericsoftware.kryo.KryoException: Unable to find class: myclass。

【问题答复】

可能性比较多，可能jar自己有问题，也有可能是因为姿势问题没有发送到executor上导致报错

8.10.5. UDF调用报错（二）


```

at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(TwoDArrayAvg.java:82)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:337)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.fun$wrapper$1(SelectOperator.scala:502)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.apply(SelectOperator.scala:521)
at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
at io.transwarp.nucleon.util.CompletionIterator.next(CompletionIterator.scala:27)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.next(CruXIteratorFactory.scala:76)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.foreach(CruXIteratorFactory.scala:69)
at io.transwarp.inceptor.execution.FileSinkOperator.processPartition(FileSinkOperator.scala:301)
at io.transwarp.inceptor.execution.FileSinkOperator$$anonfun$executeProcessFilesSinkPartition$1.apply(FileSinkOperator.scala:800)
at io.transwarp.nucleon.scheduler.ResultTask.runTask(ResultTask.scala:83)
at io.transwarp.nucleon.scheduler.Task.run(Task.scala:86)
at io.transwarp.nucleon.executor.Executor$TaskRunner$$anonfun$run$1.apply$seq$step(Executor.scala:498)
at io.transwarp.nucleon.executor.Executor$TaskRunner.run(Executor.scala:427)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.lang.Thread.run(Thread.java:750)
2022-12-23 11:38:32,070 INFO io.transwarp.nucleon.executor.CoarseGrainedExecutorBackend: Got assigned task 534
2022-12-23 11:38:32,069 ERROR io.transwarp.nucleon.executor.UncaughtExceptionHandler: Uncaught exception in thread Thread[Executor task launch worker-26,5,main],lang.IncompatibleClassChangeError: Found class org.apache.hadoop.hive.ql.udf.generic.GenericUDF$DeferredObject, but interface was expected
at com.cattl.cari.sh.cpbdd.udf.twoDArrayAvg.evaluate(TwoDArrayAvg.java:82)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:337)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:155)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.fun$wrapper$1(SelectOperator.scala:502)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.apply(SelectOperator.scala:521)
at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
at io.transwarp.nucleon.util.CompletionIterator.next(CompletionIterator.scala:27)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.next(CruXIteratorFactory.scala:76)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.foreach(CruXIteratorFactory.scala:69)
at io.transwarp.inceptor.execution.FileSinkOperator.processPartition(FileSinkOperator.scala:301)
at io.transwarp.inceptor.execution.FileSinkOperator$$anonfun$executeProcessFilesSinkPartition$1.apply(FileSinkOperator.scala:800)
at io.transwarp.nucleon.scheduler.ResultTask.runTask(ResultTask.scala:83)
at io.transwarp.nucleon.scheduler.Task.run(Task.scala:86)
at io.transwarp.nucleon.executor.Executor$TaskRunner$$anonfun$run$1.apply$seq$step(Executor.scala:498)
at io.transwarp.nucleon.executor.Executor$TaskRunner.run(Executor.scala:427)
at io.transwarp.nucleon.executor.Executor$TaskRunner.run(Executor.scala:427)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1149)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:624)
at java.util.concurrent.Thread.run(Thread.java:750)
2022-12-23 11:38:32,069 ERROR io.transwarp.nucleon.executor.UncaughtExceptionHandler: Uncaught exception in thread Thread[Executor task launch worker-20,5,main],lang.IncompatibleClassChangeError: Found class org.apache.hadoop.hive.ql.udf.generic.GenericUDF$DeferredObject, but interface was expected
at com.cattl.cari.sh.cpbdd.udf.twoDArrayAvg.evaluate(TwoDArrayAvg.java:82)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:337)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:155)
at org.apache.hadoop.hive.ql.exec.ExprNodeGenericFuncEvaluator.evaluate(ExprNodeGenericFuncEvaluator.java:130)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.fun$wrapper$1(SelectOperator.scala:502)
at io.transwarp.inceptor.execution.SelectOperator$$anonfun$2.apply(SelectOperator.scala:521)
at scala.collection.Iterator$$anon$11.next(Iterator.scala:328)
at io.transwarp.nucleon.util.CompletionIterator.next(CompletionIterator.scala:27)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.next(CruXIteratorFactory.scala:76)
at io.transwarp.inceptor.execution.cruX.CruXIteratorFactory$$anon$1.foreach(CruXIteratorFactory.scala:69)
at io.transwarp.inceptor.execution.FileSinkOperator.processPartition(FileSinkOperator.scala:301)

```

```

5 usages | override
public String[] getRequiredJars() { return null; }

5 usages | 2 overrides
public String[] getRequiredFiles() { return null; }

Evaluate the GenericUDF with the arguments.
Params: arguments – The arguments as DeferredObject, use DeferredObject.get() to get the actual
initialize call.
Returns: The

122 implementations
public abstract Object evaluate(DeferredObject[] arguments)
throws HiveException;

Get the String to be displayed in explain.

117 implementations
public abstract String getDisplayString(String[] children);

Close GenericUDF. This is only called in runtime of MapRedTask.
@Override
public void close() throws IOException {
}

```

实现的是genericUDF里面这个evaluate方法，udtf可以跑通，udf就报错这个了

【解决方法】

我们和开源的基类不一样，一个是接口，一个是抽象类，所以编译后的class不一样，运行时会上报。需要依赖社区版image里边的inceptor-exec包，然后重新编译一下。

8.10.6. 星环UDTF函数的输出相关的问题

8.10.6.1. 当函数的输出数据量特别大，应如何优化呢？

【用户问题】

请教一个问题，就是星环UDTF函数的输出数据量特别大，有没有什么优化方式，我看forward是不是只能支持一行一行的输出，不能批量输出

【问题答复】

forward确实只能一行一行输出，一般来说一个sql的性能瓶颈不是在forward这边，可能是在udtf以外的逻辑，或者是udtf本身的处理逻辑，forward其实开销不大

8.10.6.2. udtf里面能加log4j日志吗？

【用户问题】

那udtf里面能加log4j日志吗，我加了但是在/var/log/quark1目录看不到

【问题答复】

可以的，在executor节点上的这个目录的xxx executor log private static final Log LOG = LogFactory.getLog("io.transwarp.UDFTest"); LOG.info("xxxx")

8.11. 数据倾斜引起的问题

8.11.1. 星环hive表的数据倾斜导致查询跑不动了怎么解决

【用户问题】

星环orc表，由于操作不当导致数据都挤在一个分区了，现在查询任务已跑不动了，有没有什么解决办法，求解？

【问题答复】

您好，首先您需要调整下您的分区，您可以参考手册上分区表的建立方法建一个新表，分区后，将旧表的数据导入到新表，然后验证分区效果。

— 开启数据动态写入，即写入时基于分区键的值自动将数据放置到对应分区中

```
set hive.exec.dynamic.partition=true;  
set stargate.dynamic.partition.enabled=true;
```

这个也需要在导入数据前开一下。

顺便，设计适当的表分区策略是确保分区表性能最大化的关键一步，以下是一些步骤和考虑因素，可帮助您制定有效的分区策略：

- 分析查询需求：分析查询需求，重点关注经常被查询的数据的过滤条件，以选择适当的分区键，使得满足这些过滤条件的数据能够集中在同一分区中，从而优化查询性能。

- 确认数据类型：推荐使用 STRING 或时间类型的列作为分区键，通常可以帮助在数据均衡和分区数量上取得较好的平衡。
- 权衡分区规模：常规情况下，单个分区的数据量控制在 500GB 内，如果集群的 CPU 核数较多，可适当提升，此外，我们还需要关注数据的增长趋势，例如数据按时间增长，时间则是一个优秀的分区键，查询按时间范围过滤时会更高效。
- 选择分区策略：比如ArgoDB 支持范围分区和单值分区

分区类型	说明
范围分区	按照分区键的值范围来划分分区，执行分区时可基于列值分布均衡度和查询需求来自由划分范围，可避免分区间的数据规模差距过大，提升查询效率。
单值分区	将拥有相同分区键值的记录划分在同一分区中，适用于列基数较少（例如城市名）且分布较为均衡的场景。

8.12. 其他产品相关问题

8.12.1. Argodb分页能否用limit?

【问题答复】

支持这种分页查询 `select * from 表名 limit start,count`，不支持 `offset` 方式

8.12.2. ArgoDB支持对字符串进行SM3加密吗，有支持的函数吗？

【用户问题】

ArgoDB是否支持对字符串进行国密算法SM3加密？

【问题答复】

ArgoDB支持SM3国密算法，加密函数作为udf jar包可联系售后人员提供。

8.12.3. 请问argodb如果获得表的索引信息？

【用户问题】

有没有sql语句这种，获得索引名，索引字段之类的？

【问题答复】

可以通过下述 SQL 语句来查询表的详细信息，包含索引信息

```
desc formatted 表名
```

还可以通过 `select holo_index('库名', '表名');`

8.12.4. Argodb中，取最新8位日期时，REPLACE和TO_CHAR的返回结果存在BUG，求解？

【用户问题】

A	B
20230731	1
20230731	2

表TABLES，A列和B列，两行数据，最新数据日期=20230801

```
SELECT
SUM(CASE WHEN A = REPLACE(DATE_SUB(CURRENT_DATE(),1),'-','') THEN B END) 值1
,SUM(CASE WHEN A = TO_CHAR(DATE_SUB(CURRENT_DATE(),1),'YYYYMMDD') THEN B END) 值2
FROM TABLES
```

结果:
值1=0, 值2=3

```
SELECT
CASE WHEN A = REPLACE(DATE_SUB(CURRENT_DATE(),1),'-','') THEN B END 值1
,CASE WHEN A = TO_CHAR(DATE_SUB(CURRENT_DATE(),1),'YYYYMMDD') THEN B END 值2
FROM TABLES
```

结果:
值1 值2
1 1
2 2

【问题答复】

您好，请问当前 ArgoDB 使用的版本是什么，是否有使用其他方言？这里尝试使用 5.2 版本验证如下，可保持相同版本验证一下。

另外，关于函数的更多介绍，可参考 ArgoDB 开发者指南：<https://docs.transwarp.cn/#/documents-support/docs-detail/document/ARGODB-PLATFORM/5.2/010ArgoDB-Developer-Guide>

```
SELECT * FROM testdate;
```

a	b
2023-07-31	1
2023-07-31	2

查询 1 返回结果为：

值1	值2
3	3

查询 2 返回结果为

值1	值2
1	1
2	2

补充一下，表的 a 列是 string 结果也相同

```
select * from testdate;
```

a	b
20230731	1
20230731	2

8.12.5. ArgoDB创建好的存储过程的创建语句怎么查?

【问题答复】

建表语句DESC PLSQL FUNCTION EXTENDED 后面跟存储过程名称，下图示例中是 hello_world

```
0: jdbc:hive2://localhost:10000/default> DESC PLSQL FUNCTION EXTENDED hello_world;
+-----+
|          description          |
+-----+
| Prototype:                   |
| PROCEDURE gc.hello_world()   |
| Text:                        |
| CREATE OR REPLACE PROCEDURE  |
| -- 创建一个不带参数的过程 hello_world. |
| hello_world()                |
+-----+
```

8.12.6. ArgoDB 写入 HoloDesk 表报错，超过 2500 万之后数据无法写入

【用户问题】

建表语句:

```
CREATE TABLE `argo_bigdata` (
  `id` bigint NOT NULL,
  `col_num` bigint DEFAULT NULL,
  `col_var` varchar(8000) DEFAULT NULL,
  `col_bool` int DEFAULT NULL,
  `col_date` string DEFAULT NULL,
  `col_time` timestamp NOT NULL,
  `comm` varchar(1020) DEFAULT NULL
)
CLUSTERED BY (
  `id`)
INTO 10 BUCKETS
ROW FORMAT SERDE
  'io.transwarp.inceptor.memstore2.LazySimpleSerDeWrapper'
STORED BY
  'io.transwarp.inceptor.memstore2.HolodeskStorageHandler'
WITH SERDEPROPERTIES (
  'serialization.format'='1')
TBLPROPERTIES (
  'columnar.store'='true',
  'holodesk.storage.format'='normal',
  'holodesk.databasename'='argo_test',
  'holodesk.tablename'='argo_test.argo_bigdata_116cd83c-e9c9-4776-a577-865977df4143',
  'holodesk.columns.mapping'='{\"id\":\"id\",\"col_num\":\"col_num\",\"col_var\":\"col_var\",\"col_bool\":\"col_bool\",\"col_date\":\"col_date\",\"col_time\":\"col_time\",\"comm\":\"comm\"}',
  'holodesk.rowkey'='id',
  'holodesk.storage.policy'='COLD',
  'transactional'='true')
27 rows selected (0.105 seconds)
```

报错信息:

```
0: jdbc:hive2://t5:10000/default>
0: jdbc:hive2://t5:10000/default> insert into argo_bigdata(id, col_num, col_time) select 1, 990, '2022-02-09 08:40:50.0';
Error: java.lang.RuntimeException: Commit transaction: 105677 failed, due to error: java.lang.RuntimeException: Commit transaction 105677 failed, error: commit transaction found failed, status:kCommitFailedBulkTransaction (state=-,code=0)
0: jdbc:hive2://t5:10000/default>
0: jdbc:hive2://t5:10000/default>
```

数据总量:

```

27 rows selected (0.105 seconds)
0: jdbc:hive2://t5:10000/default> select count(1) from argo_bigdata;
+-----+
|  _c0  |
+-----+
| 25000000 |
+-----+
1 row selected (30.844 seconds)
0: jdbc:hive2://t5:10000/default> █

```

ArgoDB 是有什么限制吗？是不是需要设置什么参数才能写入 2500 万以上的数据呢？

【问题答复】

您好，该提示表示事务提交到存储层报错，需要确认一下 TDDMS 上对应的事务日志，分析具体写入失败的原因，您也可以联系技术支持协助排查。另外 ArgoDB 在建表时，可以不指定分桶数，默认分桶数是比较大的质数。

更多介绍，参阅 [ArgoDB 开发者指南](#)

8.12.7. 星环哪个组件支持把 Kafka 的数据直接落到 ArgoDB?

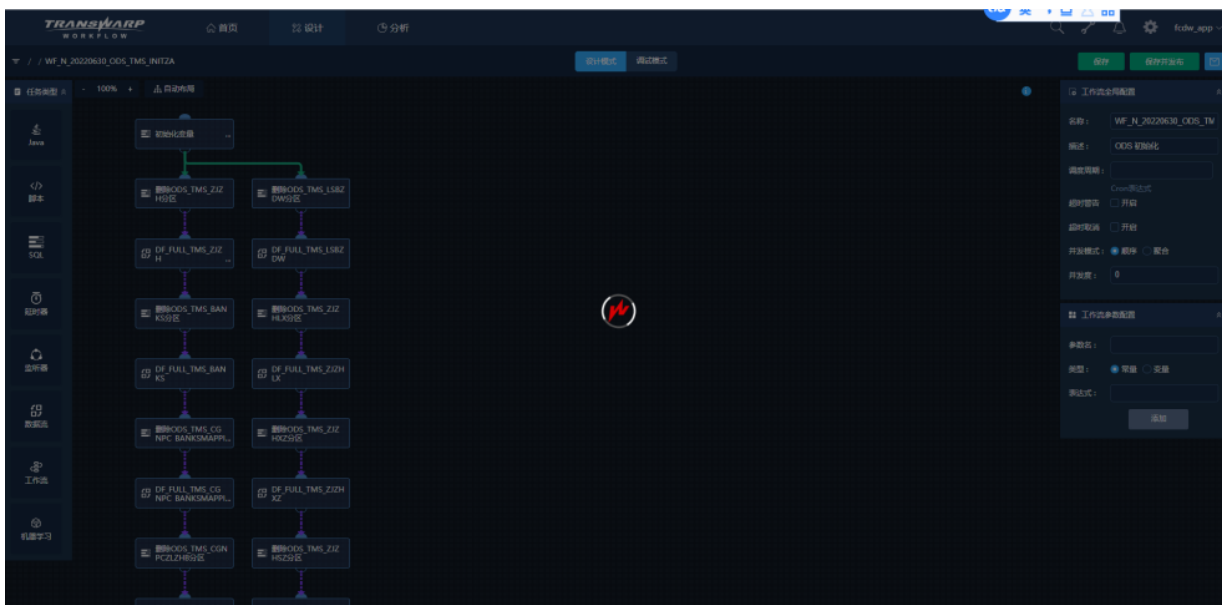
【问题答复】

ArgoDB 高性能版本在存储上做了比较多的改进，可以用 slipstream 把数据直接落库

8.12.8. Workflow 任务打不开

【用户问题】

这个任务一直打开一直转圈、但在别的电脑能打开。



【问题答复】

可以通过 F12 看看有没有什么报错，包括 console 也看一下。可能是因为任务流里面任务比较多，可以检查一下浏览器，如果任务较多的话，也可能是前端渲染的问题。

【用户结论】

我们使用的是谷歌浏览器，任务是较多的。然后我们换了别的浏览器登录解决了这个问题。还有 F12 模式下切换手机电脑显示模式偶尔也可以解决这个问题、谢谢！

8.12.9. Workflow内置时间参数有哪些？

【用户问题】

workflow如何调用内置的时间参数，具体的时间参数都有哪些，表达式中如何调用？

【问题答复】

举例：

```
schedule.date(): 返回调度的日期，精确到天；
schedule.datetime(): 返回调度时间，精确到秒；
schedule.conciseDate(): 返回调度日期；
schedule.conciseDatetime(): 返回调度时间；
```

用法示例：

```
schedule.plusYears(1).date(): 返回调度日期向后延迟一年之后的日期。假如调度日期是2017-08-17，那么返回2018-08-17；
schedule.plusYears(-1).datetime(): 返回调度时间提前一年的时间。假如调度日期是2017-08-17 11:04:25，那么返回2016-08-17 11:04:25；
schedule.plusMonths(1).datetime(): 将调度时间向后延迟一个月；
schedule.plusWeeks(4).datetime(): 将调度时间向后延迟4周；
schedule.plusDays(30).datetime(): 将调度时间向后延迟30天；
```

还有一些注意事项以及其他参数等等，具体如何调用可以联系售后获取TDS操作手册哦~

8.12.10. Sophon添加了quark数据集连接之后，会显示无状态请更新

【问题答复】

会话未启动，可以重启会话试试。

【问题后续】

用户重启后依然是一样的报错信息，且无论是上传的数据集还是其他数据集都无法预览，在实验中也无法读取。

随后用户尝试上传数据，显示会话没有启动或准备好，用户需要在社区版manager页面重启下服务，重启顺序，session→data→vlab。用户页面重启后依然不行，那么在这个情况下用户则需要查看sophon-session的pod的日志。

查看日志后发现线程卡在jar包上传了，因此有可能是环境问题，在启动的过程中卡住了。

用户需要看下session服务里jvm进程的jstack，看看间隔一定时间打几次jstack，有没有线程卡在什么地方进一步排查。

8.12.11. StellarDB v4.1版本，有详细的python Sdk 说明demo么？

【用户问题】

StellarDB v4.1版本，有详细的python Sdk 说明demo么？能够支持查询，建图，创建点边数据的。求发一份，目前能够看到的就是查询demo，有创建 编辑的demo么？

【问题答复】

示例：

```
conn = stellar_hive.StellarConnection(host="host1", port=10000, graph_name='family')
cur = conn.cursor()
cur.execute('config query.lang cypher')
cur.execute('use graph family')
```

这里的use graph xxx直接换成建图语句就可以了
比如换成:
create graph family with schema (:vertex) [:edge] graphproperties:{`graph.shard.number`:3};

8.12.12. 执行Spark任务，executor节点报错显示UnknownHost

【用户问题】

将Spark任务从非TDH集群节点提交到TDH集群上去执行，任务跑起来之后在TDH集群上也启动相关的任务执行，但是执行的executor节点(TDH集群节点)会报UnknownHost错误。不能识别到我提交任务的机器。

请问下我们提交任务的机器hostname和ip映射信息是否需要配置到TDH集群对应的k8s集群中？为什么配置了/etc/hosts/文件，把这提交任务的节点hostname与ip映射添加到文件中还是无法解决？

【核心原因】

/etc/hosts文件并非容器挂载文件，因此修改/etc/hosts文件，并不会影响容器里面的内容

【问题答复】

需要手动修改集群三个节点下的挂载文件（/etc/transwarp/conf/hosts），分别添加提交任务的这台非集群节点的hostname与ip映射。添加完成后在Manager界面重启一下Yarn服务，再试一下任务是否可以执行。

8.12.13. TDH上有 kafka的管理界面吗？

【用户问题】 TDH上有 kafka的管理界面吗？例如 创建 topic ,查看topic 信息等。或者是否有推荐的web 组件，方便查看 topic 信息？

【问题答复】

TDH上kafka的管理界面可以管理服务的启停迁移，topic相关信息还是要通过kafka命令行去查看或者API。推荐使用星环高吞吐分布式NoSQL数据库Event Store替代开源kafka，可以帮助您快速开发日志收集、应用监控、流式数据处理、在线分析等应用。

客户服务

技术支持

感谢您使用星环信息科技（上海）股份有限公司的产品和服务。如您在产品使用或服务中有任何技术问题，可以通过星环开发者论坛途径找到我们的技术人员给予解答。

星环开发者论坛支持：<https://community.transwarp.cn/>

email: support@transwarp.io

官方网址：<http://www.transwarp.cn/>

意见反馈

如果您在系统安装，配置和使用中发现任何产品问题，可以通过以下方式反馈：

email: support@transwarp.io

感谢您的支持和反馈，我们一直在努力！