TCC简介

(仅供内部使用）

For internal use only

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 拟制:  Prepared by |  |  | 日期：  Date |  |
| 审核:  Reviewed by |  |  | 日期：  Date | yyyy-mm-dd |
| 批准:  Granted by |  |  | 日期：  Date | yyyy-mm-dd |



华为技术有限公司

Huawei Technologies Co., Ltd.

版权所有 侵权必究

All rights reserved

[1 TCC介绍 4](#_Toc434241289)

[1.1 TCC功能特性 4](#_Toc434241290)

[1.1.1 调度模块 4](#_Toc434241291)

[1.1.2 执行模块 5](#_Toc434241292)

[1.1.3 管理模块 6](#_Toc434241293)

[2 代码 8](#_Toc434241294)

[2.1 配置库路径 8](#_Toc434241295)

[2.2 工程框架 8](#_Toc434241296)

[2.2.1 使用组件 8](#_Toc434241297)

[2.2.2 初始化 8](#_Toc434241298)

[2.3 目录结构 10](#_Toc434241299)

[2.3.1 工程目录 10](#_Toc434241300)

[2.3.2 源码目录 11](#_Toc434241301)

[2.4 核心模块逻辑 12](#_Toc434241302)

[2.4.1 调度模块：TccServiceImpl 类的scheduleTask方法 12](#_Toc434241303)

[2.4.2 执行模块：CycleTask类的run方法 12](#_Toc434241304)

[2.4.3 管理模块：以OS组管理功能为例说明需要涉及到的文件 12](#_Toc434241305)

[3 现网维护 13](#_Toc434241306)

[3.1 负责业务管理、节点管理、权限管理、TCC配置（具体可参考《TCC操作使用手册.docx》） 13](#_Toc434241307)

[3.2 TCC环境维护 14](#_Toc434241308)

[3.2.1 数据库配置 14](#_Toc434241309)

[3.2.2 TCC模块配置 14](#_Toc434241310)

[3.2.3 日志文件配置 16](#_Toc434241311)

[3.2.4 SUSE远程登陆配置 17](#_Toc434241312)

[3.2.5 Sudo配置 17](#_Toc434241313)

[3.2.6 脚本部署 18](#_Toc434241314)

[**4** **FAQ** 18](#_Toc434241315)

[4.1 TCC如何调度任务周期？ 18](#_Toc434241316)

[4.2 如何判断任务周期的依赖关系已经满足？ 18](#_Toc434241317)

[4.3 TCC如何运行任务周期？ 19](#_Toc434241318)

[4.4 如何定位任务周期没有运行的问题？ 19](#_Toc434241319)

[4.5 任务周期的正向依赖树和反向依赖树有什么作用？ 20](#_Toc434241320)

[4.6 任务周期是错误的运行状态，如何定位出错原因？ 20](#_Toc434241321)

[4.7 多批次任务对传递过来的文件有哪些约束？ 20](#_Toc434241322)

[4.8 任务步骤的执行命令有哪些约束？ 20](#_Toc434241323)

[4.9 为什么修改了任务或者任务步骤的配置对运行的任务周期没有立即生效？ 21](#_Toc434241324)

[4.10 为什么在周期管理页面的【源任务周期集合Tab页】看到的任务周期状态和在其它页面看到的有些不一致？ 21](#_Toc434241325)

[4.11 为什么周期管理的【待处理任务周期集合Tab页】和批量重做的【待处理任务集合Tab页】都有重做功能？ 21](#_Toc434241326)

[**5** **其它材料** 21](#_Toc434241327)

# TCC介绍

任务控制中心是BI系统的调度控制中心，按照依赖关系、优先级等周期性执行任务的所有步骤脚本，从而保证hadoop数据仓库中的所有统计和分析过程得以正确而有效的执行。包括调度模块、执行模块、TCC管理模块。

**调度模块**负责选择依赖关系满足并且时间上可运行的任务进入等待队列、然后选择优先级高的任务进入运行队列。

**执行模块**负责按顺序登陆指定的主机执行步骤脚本、实时取回执行日志、超时控制、执行出错重试、反馈执行状态、持久化步骤和任务周期的执行状态以及中途停止正在执行的任务。

**TCC管理模块**包括任务管理、任务步骤管理、周期管理、集群管理、节点管理、业务管理、权限控制、告警、审计、任务批量重做、任务的运行状态监控等等

## TCC功能特性

### 调度模块

1. 为不同周期类型（年、月、周、天、时）的任务创建相应的任务周期，并且周期长度可以指定；

比如：任务A的周期类型为天，并且周期长度为2，那么TCC会每2天为A任务创建一个任务周期，周期标识为周期的起始时间，格式固定为YYYYMMDD-00;

1. 不同的任务可以配置不同的起始时间，只有落在起始时间之后任务周期才会被调度；

比如：天类型任务A的起始时间2012-10-11 00::00:10，那么最小的任务周期Id就是

20121012-00;

1. 任务周期为最小的调度单元，任务周期的优先级、依赖关系、周期偏移都来源于任务中的配置，TCC调度模块按照任务周期的依赖关系是否满足、周期偏移剩余时间是否为0和优先级大小以及权重来决定是否运行；

依赖关系：大周期类型任务依赖小周期类型任务时，如果是全周期依赖，当前任务周期依赖于它所包含的所有小任务周期，否则，当前任务周期依赖于它所包含的所有小任务周期的最后一个任务周期；小周期类型任务依赖大周期类型任务时，当前任务周期依赖于包含它的大任务周期；相同周期类型任务依赖时，当前任务周期依赖于另一任务的相同任务周期；

比如：天类型任务A依赖于小时任务B，A的当前任务周期c为20121011-00，那么c就依赖于任务B的20121011-23，如果还指定全周期依赖，那么c就依赖于任务B的20121011-00到20121011-23的所有周期；

比如：天类型任务A依赖于天任务B，A的当前任务周期c为20121011-00，那么c就依赖于任务B的20121011-00

比如：小时类型任务A依赖于天任务B，A的当前任务周期c为20121011-01，那么c就依赖于任务B的20121011-00

依赖关系是否满足：当任务A的当前周期c所依赖的任务周期的运行状态全部为成功时，则认为c的依赖关系已经满足；如果A依赖的某个任务B已经指定为忽略执行出错，那么B的所有任务周期中执行出错的也当做是成功。

周期偏移剩余时间：MAX(0, 周期ID+周期偏移-系统时间)；

优先级大小：优先级数值越小的优先级越高；

***注意***：依赖关系已经满足、周期偏移剩余时间为0、优先级最高的所有任务周期中权重刚好小于等于剩余的并发运行数的任务周期越先运行；

1. 当历史任务周期的运行状态被重新改成初始化后，TCC调度程序会重新正常调度；
2. 为保证系统资源不被过度占用，TCC调度模块保证并行执行的任务周期数不能超过配置的最大周期数；
3. 任务支持权重W，表示该任务的一个任务周期运行时占用的并发周期数为M而不是1。

如果任务比较耗系统资源，就可以将权重配置大一些，反之，配置小一些；

1. TCC可能中途停止，重新启动TCC后支持重新调度原来未执行或正执行的任务周期，支持任务周期从上次停止的步骤处继续执行；
2. 可以停止或启动任务；

停止任务后TCC将不再调度该任务的周期，并停止正在运行的所有周期。

1. 可以启动或停止任务步骤；

### 执行模块

1. 支持ssh协议利用信任关系认证登录指定主机指定用户，执行指定命令脚本；
2. 支持出错重试。命令执行失败后，如果配置了多个节点，可以从一个节点切换到下一个节点重新执行。最后一个节点执行失败后，会多次延时重新执行命令；
3. 支持超时停止。命令执行超过指定时间后，自动停止，释放资源；
4. 记录执行日志。命令运行的日志会记录下来，方便出错后定位；
5. 当任务周期开始执行时，创建批次、步骤，并按照顺序执行每个批次的所有步骤，执行的状态更新到相应的运行状态表中。

创建批次时分两种情况：对于批次任务，根据文件结束标识以不同的方式来创建批次运行状态（一个批次对应一个文件）；对于非批次任务，默认创建一个批次运行状态。

步骤运行状态：如果步骤执行出错，则重试，在重试次数范围内执行成功，则步骤的运行状态为成功，否则任务步骤运行状态为出错；如果步骤的执行时间（包括重试）超过超时时间，任务步骤运行状态为超时；

批次运行状态：如果批次的所有步骤运行状态中有一个是超时，则批次的运行状态为超时；否则，如果有一个是出错，则批次的运行状态为出错；上述情况之外的情况，则批次的运行状态为成功；

任务（周期）运行状态：如果是批任务，没有一个文件批次的话，任务运行状态为文件未到达；否则，如果所有批次的运行状态中有一个是超时，则任务的运行状态为超时；否则，如果有一个是出错，则任务的运行状态为出错；上述情况之外的情况，则任务的运行状态为成功；

1. 支持集成重做，仅对非批次、小时或天周期类型、顺序依赖的任务有效；

集成重做：如果任务配置了集成重做开始与结束时间以及集成重做天数N，那么就将每N天内的历史任务周期作为一个任务周期执行（TCC实际执行N天内的最后一个任务周期，其它的任务周期的均虚拟执行）；

### 管理模块

1. 业务管理：业务的增、删、改、查。对于每个业务，可配置业务名、业务Id、业务联系人、告警邮箱列表、告警手机号码列表（暂不支持）。
2. 集群管理：节点的增、删、改、查。对于每个集群，对集群内的资源（并发数、节点内并发数等）进行单独控制；
3. 节点管理：节点的增、删、改、查。对于每个部署节点，可配置节点名、主机、端口等信息。
4. OS组管理：操作系统组的增、删、改、查。对于每个OS组，可以设置组名已经绑定的业务。这些OS组需要事先在linux操作系统中创建，然后在改模块中录入。
5. OS用户管理：操作系统用户的增、删、改、查。对于每个OS用户，可以设置用户名和密码。这些OS用户需要事先在linux操作系统中创建，然后在改模块中录入。
6. 角色管理：角色的增、删、改、查。对于每个角色，可以设置角色名以及为角色定义不同业务的不同任务组的权限。权限包括4种：完全权限、执行权限、查询权限、无权限。
7. 用户管理：用户的增、删、改、查。对于每个用户，可以绑定角色。
8. 操作审计：可以审查操作员在不同时间范围内进行的各种操作。通过组合业务类型、任务名称、操作类型、操作员、时间范围来查询操作员的操作记录。
9. 日志查询：组合操作员、日志记录、时间范围来查询TCC产生的日志，用于快速定位TCC故障或者任务执行出错等问题原因。
10. TCC配置：集中配置TCC需要的各种参数，主要包括基准时间、最大并发数、是否发送提醒邮件等等。
11. 任务管理：任务和任务步骤的创建、修改、删除、查看、复制新增、导入、导出、启动、停止以及任务的条件查询和任务的运行状态查看、任务告警配置；
12. 周期管理：任务周期的条件查询、任务周期的重做、任务周期的整棵依赖树（近期少量）的重做、任务周期的详情查看、任务正反向依赖树的展示、端到端任务时间统计、任务周期的集成重做等等；
13. 批量重做：实现指定批量任务的指定时间范围的任务周期重做。
14. 告警管理：TCC根据任务的告警配置产生相应的告警记录，通过组合业务类型、告警类型、告警级别、告警状态、时间范围类筛选要处理的告警记录。可为每条告警记录问题原因，解决措施。

告警类型包括：

1. 任务结束时返回异常（严重）
2. 任务失败（严重）
3. 任务发生重做 (一般)
4. 任务执行时间超时（严重）
5. 任务到达最迟启动时间时仍未启动（严重）
6. 任务到达最迟结束时间时仍未结束（严重）
7. 弱依赖任务未完成时强行启动任务(一般)
8. 周期队列：查看处于等待队列或者运行队列中的任务周期。只有依赖关系满足和周期偏移剩余时间（周期偏移剩余时间=MAX(0,周期ID+周期偏移-系统时间)）为0的任务周期才会被TCC调度到等待队列中，优先级高的任务周期优先从等待队列调度进运行队列运行。

# 代码

## 配置库路径

https://119.145.15.45:3690/svn/TC\_EMUI\_DataService\_SVN/合作方目录/平台工具/code/TaskControlCenter\_RS

## 工程框架

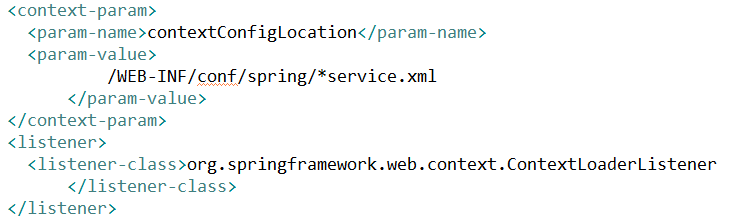
### 使用组件

spring框架+struts+mybatis+mysql+tomcat+jdk+eclipse+navicate lite

### 初始化

Web项目的入口文件是web.xml

1. web.xml是由web容器（如tomcat）解析，其中通过配置Listener拉起spring框架和struts框架





1. spring框架读取/WEB-INF/conf/spring/目录下所有以service.xml结尾的文件，按照依赖关系初始化bean，不同文件中的bean提供不同服务

tcc.common.config.init.service.xml：处理tcc配置文件，加载tcc的配置文件

tcc.common.mail.service.xml：邮件服务，stmp地址、账号等

tcc.ds.service.xml：数据源，加载tcc.jdbc.properties等jdbc配置，连接数据库

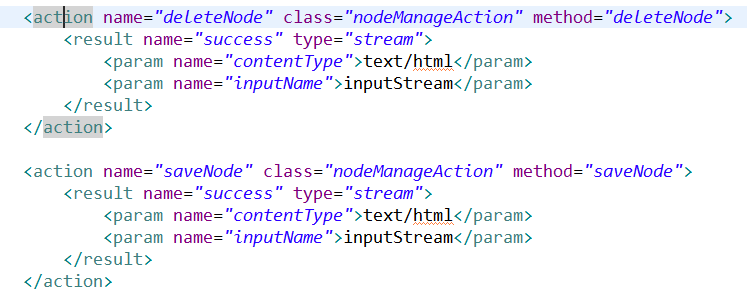
tcc.mybatis.service.xml：mybatis服务，将sql语句与业务实现分离

tcc.service.xml：业务层实例，包括所有管理功能的服务，如：业务管理、节点管理等

tcc.threadpool.service.xml：线程池服务，调用任务周期的run方法

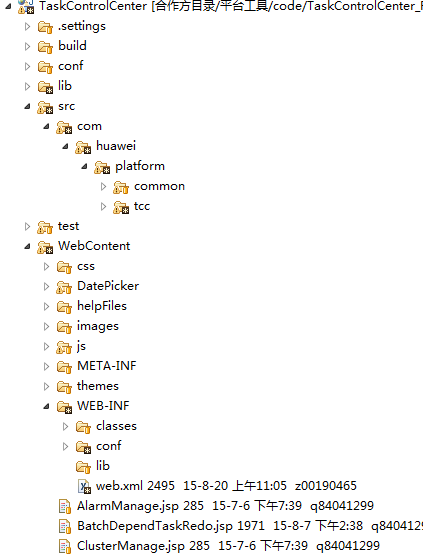
tcc.timetask.service.xml：定时器服务，用于定时调用调度模块入口函数（scheduleTask）、邮件发送函数（*sendEmails*）、ssh通道检查函数（*checkAllChannels*）

1. struts框架可以方便的将http请求地址转发到指定对象的指定方法中（通过配置struts.xml文件的action），不用每个请求都实现一个servlet，非常方便



## 目录结构

### 工程目录



build目录：构建TCC版本包的脚本

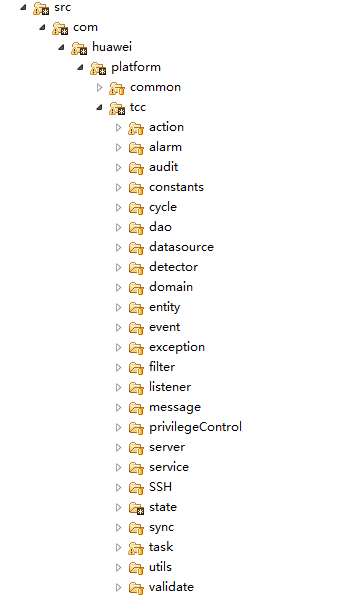
conf：TCC的配置文件目录

lib:外部库目录

src:源码目录

WebContent：存放web.xml文件，jsp，css，js等资源文件

### 源码目录



action:action类，管理功能请求实现

alarm:告警模块

audit：审计模块

dao：数据库层

service:业务层，调度逻辑

ssh:ssh远程执行模块

## 核心模块逻辑

### 调度模块：TccServiceImpl 类的scheduleTask方法

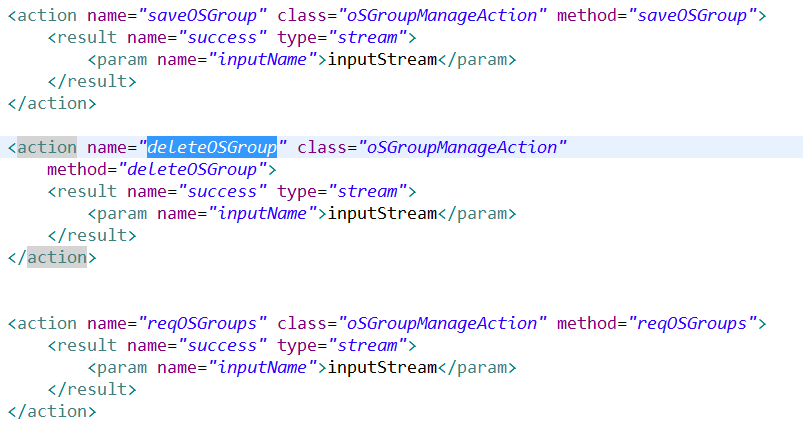


### 执行模块：CycleTask类的run方法

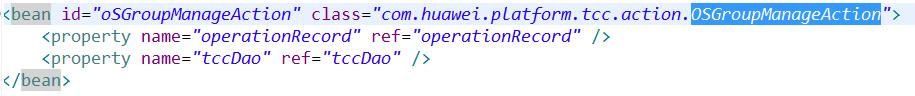


### 管理模块：以OS组管理功能为例说明需要涉及到的文件

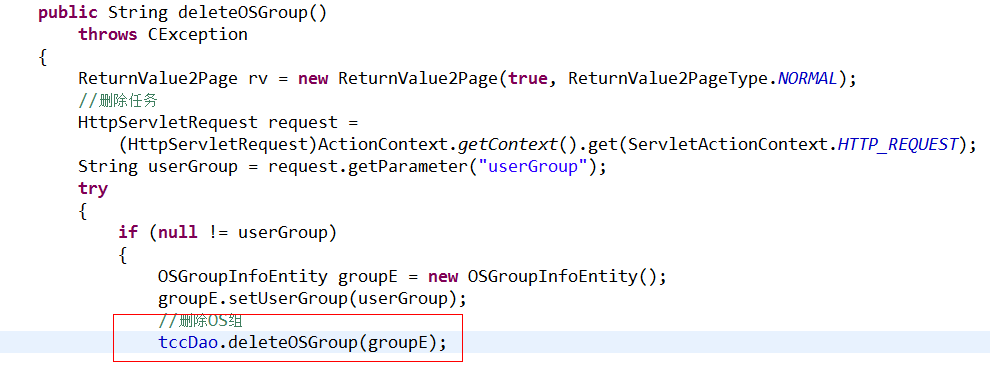
OSGroupManage.jsp页面中包含：os组的增、删、改、查功能，对应的http请求为saveOSGroup、deleteOSGroup、saveOSGroup、reqOSGroups，对应struts.xml中的配置为



*oSGroupManageAction对象配置在tcc.service.xml文件中*



四个请求的实现在*com.huawei.platform.tcc.action.OSGroupManageAction类中，分别对应*saveOSGroup、deleteOSGroup、saveOSGroup、reqOSGroups方法，其中deleteOSGroup方法调用了dao层的deleteOSGroup方法

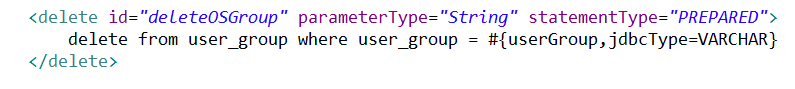


tccDao.deleteOSGroup方法通过“com.huawei.platform.tcc.dao.OSGroupInfo.deleteOSGroup”键从OSGroupInfoMapper.xml文件中找到对应的sql语句

TccDaoImpl.java文件：



OSGroupInfoMapper.xml文件：



# 现网维护

## 负责[业务管理](https://emui.huawei.com/d/TaskControlCenter/ServiceManage.jsp)、[节点管理](https://emui.huawei.com/d/TaskControlCenter/NodeManage.jsp)、[权限管理](javascript:menuClick(1.1,2))、TCC配置（具体可参考《TCC操作使用手册.docx》）

使用system用户登录<https://emui.huawei.com/d/TaskControlCenter/SysAdminIndex.jsp>

## TCC环境维护

申请跳板机权限：[TCCTest@10.41.24.84](mailto:TCCTest@10.41.24.84)

依赖组件：10.41.24.83上部署的邮件系统组件

部署路径：/home/TCCTest/TCCTest/TccServer/apache-tomcat-6.0.35\_safe/webapps/TaskControlCenter

jdbc配置文件路径：$TOMCAT\_HOME/webapps/TaskControlCenter/WEB-INF/classes/log4j.properties

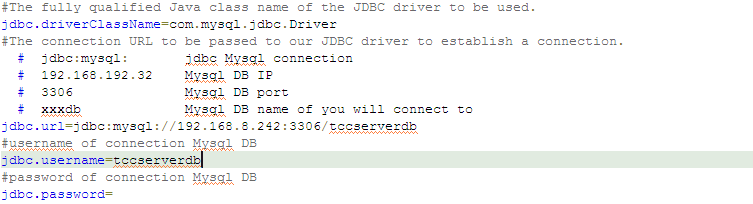
$TOMCAT\_HOME/webapps/TaskControlCenter/WEB-INF/classes/tcc.jdbc.properties

系统配置文件路径：

$TOMCAT\_HOME/webapps/TaskControlCenter/WEB-INF/conf/common/resource/systemconfig/tcc.sysconfig.properties

### 数据库配置

配置文件为TaskControlCenter\WEB-INF\classes\tcc.jdbc.properties



应该至少修改下面三个配置项修改：

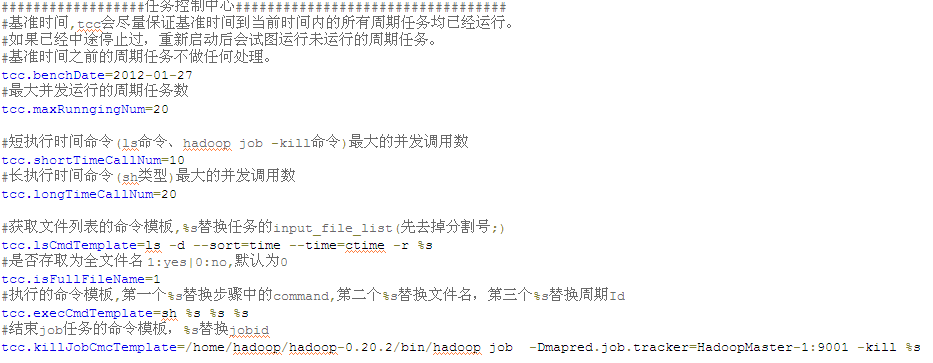
jdbc.url TCC使用的数据库地址

jdbc.username 数据库用户名

jdbc.password 数据库密码，需经过AES128加密处理

### TCC模块配置

配置文件为TaskControlCenter\WEB-INF\conf\common\resource\systemconfig\tcc.sysconfig.properties



应该至少修改下面五个配置项修改：

tcc.benchDate=***2012-01-27***

基准时间是一个非常重要的参数，设定的天范围最好是1-27（因为每个月都能保证至少有27天，如果修改成2012-01-29号的话，月类型的任务周期可能就不具备每月29号的意义了。），tcc后台将这个日期作为时间的起始标准，请勿随便调整基准时间，否则依赖关系可能不起作用；

两种情况可以调整tcc.benchDate：

1、任务运行状态表不存在任何记录时，可以随意调整tcc.benchDate；

2、任务运行状态表存在记录时，tcc.benchDate仅可以调整到表中的最大周期之后；

tcc.maxRunngingNum=***20***

同一个时间内运行并发运行的最大任务周期数，这个配置项也作为任务表的权重的最大值使用。该参数通常不应该超过tcc.longTimeCallNum，否则可能会导致任务周期执行时，调用壳程序被阻塞，反而影响实际执行时间。

tcc.shortTimeCallNum=***10***

短执行时间命令(ls命令、hadoop job -kill命令)最大的并发调用数，受限于壳程序远程登录并执行命令的主机所允许的最大登录用户数（tcc.shortTimeCallNum+ tcc.longTimeCallNum必需不超过***允许同时远程登陆的最大用户数***，修改参见2.5 SUSE远程登陆配置）；

tcc.longTimeCallNum=***20***

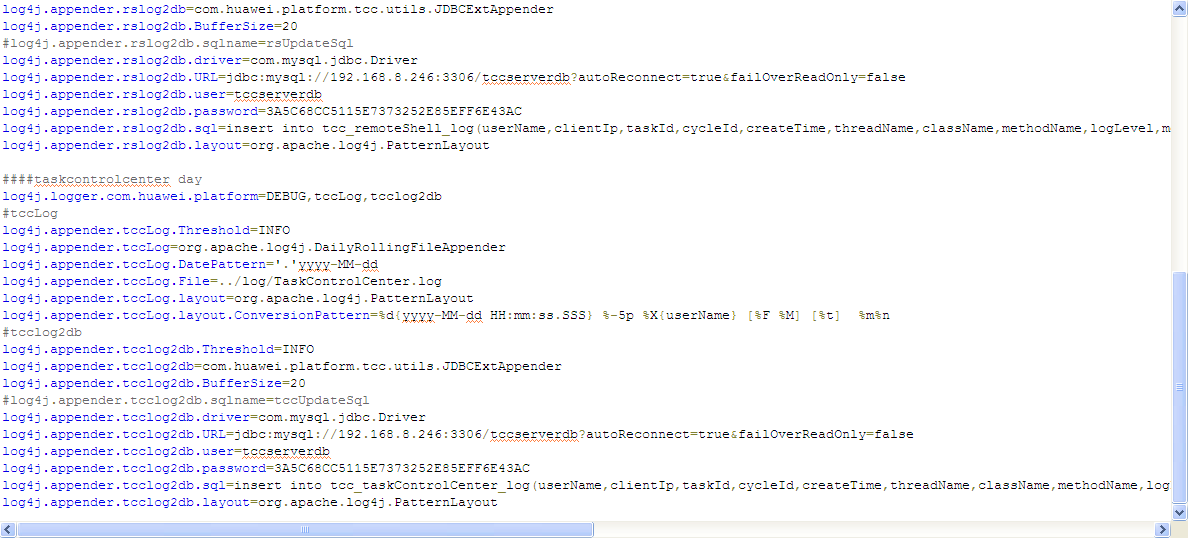
长执行时间命令(sh类型)最大的并发调用数，受限于壳程序远程登录并执行命令的主机所允许的最大登录用户数（tcc.shortTimeCallNum+ tcc.longTimeCallNum必需不超过***允许同时远程登陆的最大用户数***，修改参见2.5 SUSE远程登陆配置）；

tcc.killJobCmcTemplate=***/home/hadoop/hadoop-0.20.2/bin/***hadoop job -Dmapred.job.tracker=***HadoopMaster-1:9001*** -kill %s

结束job任务的命令模板，%s替换jobid。根据hadoop安装的路径以及tracker节点的不同进行修改

### 日志文件配置

配置文件为TaskControlCenter\WEB-INF\classes\log4j.properties



应该至少修改下面6个配置项修改：

log4j.appender.rslog2db.URL TCC使用的数据库地址（同2.1数据库配置中的jdbc.url）

log4j.appender.rslog2db.user 用户名（同3.2.3数据库配置中的jdbc.username）

log4j.appender.rslog2db.password密码（同3.2.3数据库配置中的jdbc.password）

log4j.appender.tcclog2db.URL TCC使用的数据库地址（同3.2.3数据库配置中的jdbc.url）

log4j.appender.tcclog2db.user 数据库用户名（同3.2.3数据库配置中的jdbc.username）

log4j.appender.tcclog2db.password 数据库密码（同3.2.3数据库配置中的jdbc.password）

也可以通过修改阈值的等级值来记录不同的日志级别：

log4j.appender.Console.Threshold=***INFO***

log4j.appender.rsLog.Threshold=***INFO***

log4j.appender.rslog2db.Threshold=***INFO***

log4j.appender.tccLog.Threshold=***INFO***

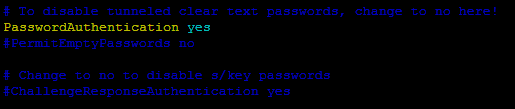
log4j.appender.tcclog2db.Threshold=***INFO***

Console为控制台输出，rsLog为远程壳输出到文件，rslog2db为远程壳输出到文件数据库，tccLog为TCC日志输出到文件，tcclog2db为TCC日志输出到数据库。

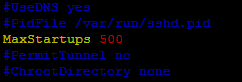
### SUSE远程登陆配置

默认的SUSE系统配置是不允许在程序中进行远程登陆认证的，所以为了使用远程壳程序需要做如下的修改：

#vi /etc/ssh/sshd\_config



将PasswordAuthentication配置项改成yes



将MaxStartups（允许同时远程登陆的最大用户数）配置项改成需要的值，通常不应太高

# service sshd restart

### Sudo配置

为了让TCC执行脚本时不输入密码，使用sudo来执行不同用户下的脚本，需要做如下的配置（hadoop为TCC部署时的用户）：

1. 使用root用户执行如下命令：

visudo -f /etc/sudoers

1. 查找下面语句的位置

# User privilege specification

root ALL=(ALL) ALL

1. 在上面的语句下面添加一条记录

hadoop localhost=(ALL) /usr/bin/sh /bin/sh

1. 查找下面语句的位置

# Same thing without a password

# %wheel ALL=(ALL) NOPASSWD: ALL

1. 在上面的语句下面添加一条记录

hadoop ALL=(ALL) NOPASSWD: ALL

1. 查找下面语句的位置

Defaults env\_reset

1. 修改上面的语句为如下

Defaults !env\_reset

### 脚本部署

killtree.sh和rebootTomcate.sh脚本需要放到$CATALINA\_HOME/bin/目录下，并

在TCC所属用户下执行如下命令赋予可执行权限：

chmod ugo+x killtree.sh

chmod u+x rebootTomcate.sh

1. **FAQ**
   1. TCC如何调度任务周期？

任务周期是最小的调度单元，任务周期的优先级、依赖关系、周期偏移都来源于任务中的配置，TCC调度模块首先会获取所有已经启用的任务，为每个任务获取周期偏移剩余时间为0（周期偏移剩余时间=MAX(0, 周期ID+周期偏移-系统时间)）而且状态为初始化或者已开始的任务周期，然后选择依赖关系已经满足的所有任务周期到等待队列中，最后从等待队列中选择优先级最高的任务周期到运行队列中并开始运行任务周期，直到运行队列满或者等待队列空。

* 1. 如何判断任务周期的依赖关系已经满足？

通过任务配置的依赖关系可以算出当前任务周期所依赖的全部任务周期，当依赖的任务周期的运行状态全部为“执行成功”时，就说当前任务周期的依赖关系已经满足。但是这里面存在特例，如果任务A依赖的某个任务B已经指定为忽略错误，那么B的所有任务周期中执行出错的也当做是“执行成功”；如果当任务A的当前周期c依赖本任务的其它“虚拟执行成功”的任务周期，也当做是“执行成功”。

* 1. TCC如何运行任务周期？

当任务周期开始运行时，创建批次、步骤，并按照顺序执行每个批次的所有步骤，执行的状态会更新到相应的运行状态表中。

创建批次时分两种情况：对于批次任务，根据文件结束标识以不同的方式来创建批次运行状态（一个批次对应一个文件）；对于非批次任务，默认创建一个批次运行状态。

步骤运行状态：如果步骤执行出错，则重试，在重试次数范围内执行成功，则步骤的运行状态为成功，否则任务步骤运行状态为出错；如果步骤的执行时间（包括重试）超过超时时间，任务步骤运行状态为超时；

批次运行状态：如果批次的所有步骤运行状态中有一个是超时，则批次的运行状态为超时；否则，如果有一个是出错，则批次的运行状态为出错；上述情况之外的情况，则批次的运行状态为成功；

任务（周期）运行状态：如果是批任务，没有一个文件批次的话，任务运行状态为文件未到达；否则，如果所有批次的运行状态中有一个是超时，则任务的运行状态为超时；否则，如果有一个是出错，则任务的运行状态为出错；上述情况之外的情况，则任务的运行状态为成功；

* 1. 如何定位任务周期没有运行的问题？

任务周期运行需要满足5个条件：

1. 任务启用；
2. 任务周期状态为初始化或已开始；
3. 依赖关系全部满足；
4. 周期偏移剩余时间为0；
5. 运行队列未满；

通过判断这5个条件，我们就可以定位出原因了。

第1步：查看指定任务周期的任务周期详细页面（参见周期管理）

通过【任务信息Tab页】判读是否满足条件①，通过【任务周期运行状态Tab页】判断是否满足条件②，通过【依赖的任务周期运行状态Tab页】判断是否满足条件③和④。

第2步：如果第1步发现的所有条件都满足了，那么可以打开周期队列页面

在【TCC运行队列Tab页】中判断是否满足条件⑤，此时如果不满足条件⑤，必然可以在【TCC等待队列Tab页】中找到该任务周期。

* 1. 任务周期的正向依赖树和反向依赖树有什么作用？

正向依赖树是由父节点任务周期依赖子节点任务周期构成，通过查看状态可以很方便的定位依赖关系是否正确以及树上的哪些任务周期间接影响了当前任务周期。

因为任务周期之间存在着依赖关系，越处于依赖关系底层的任务周期越关键，这种任务周期可能因为各种原因（数据的问题或者脚本有问题）在执行过程表现正确，但实际上是不正确的，一旦发现这种问题便需要对任务周期进行重做，而且所有直接或者间接依赖该任务周期的任务周期（即反向依赖树）都应该重做，这就是引入反向依赖树的原因。通过查看任务周期的反向依赖树可以知道当前任务周期的影响面有多广。

* 1. 任务周期是错误的运行状态，如何定位出错原因？

打开指定任务周期的任务周期详细页面（参见周期管理），通过【远程壳输出日志Tab页】查看步骤的运行日志，可以分析出原因，如果还是没有找到，可到【任务周期相关日志Tab页】中查看是否有error或者warn级别的日志，从而分析错误原因。

* 1. 多批次任务对传递过来的文件有哪些约束？

任务周期运行时，只会处理本周期内的文件，所以文件名中必需包含时间信息，即带上“***\_yyyyMMddHH***”格式的字符串。

* 1. 任务步骤的执行命令有哪些约束？

1、执行命令必需通过退出码0或者非0反馈执行结果是成功还是失败的。

2、可能因为各种原因，操作员会中途停止TCC任务，这将导致TCC停止正运行的任务周期，最终会停止正在运行的*命令进程*，为了能让TCC能停止掉*命令进程*启动的子进程树，需要*命令进程*能将自己的进程号*PID*以“REQUEST\_KILL\_PID:*PID*”格式写到输出流的独立行中*。*例如，只需要*在*/home/hadoop/HiBI/HIBI/exec-hive.sh脚本的第1条可执行命令前插入echo REQUEST\_KILL\_PID:$$即可。

* 1. 为什么修改了任务或者任务步骤的配置对运行的任务周期没有立即生效？

TCC开始运行任务周期时，就加载了任务和任务步骤的配置，所以一旦任务周期事先被调度到运行队列中运行，任务或者任务步骤配置的修改均对这些任务周期无效；所以修改任务或者任务步骤配置时，应该先**停用任务**，然后再修改配置，修改完成后再**启动任务。**

* 1. 为什么在周期管理页面的【源任务周期集合Tab页】看到的任务周期状态和在其它页面看到的有些不一致？

因为周期管理页面的【源任务周期集合Tab页】有周期查询功能，为了效率考虑，任务周期的状态是完全从数据库表中获取，而其它页面的状态是先从数据库表中获取然后经过再次运算所得，这样可以得到更加细致的状态。所以，任务周期状态可能存在差异。

* 1. 为什么周期管理的【待处理任务周期集合Tab页】和批量重做的【待处理任务集合Tab页】都有重做功能？

周期管理的重做功能，操作的是每一个任务周期，操作员可以任意筛选，适合近期的任务周期集合的重做；

批量重做的重做功能，操作的是每一个任务，操作员只能筛选任务，配置重做时间范围，但是可以方便处理大量任务周期重做的问题。

1. **其它材料**

《TCC第一轮迭代已经实现的功能列表.docx》

《TCC操作使用手册.docx》、

《TCC第二轮迭代实现方案.docx》

《TCC第三轮迭代需求.xlsx》

《TCC任务周期执行流程.pdf》

《TCC调度模块流程图.pdf》