|  |
| --- |
| 任务管理  目录  [任务管理 1](#_Toc345331343)  [1 图标说明 1](#_Toc345331344)  [2 任务查询 2](#_Toc345331345)  [3 任务名模糊查询与精确查询 2](#_Toc345331346)  [4 新建任务 3](#_Toc345331347)  [5 任务的主要配置项 4](#_Toc345331348)  [6 新建任务步骤页面 7](#_Toc345331349)  [7 任务步骤的主要配置项 8](#_Toc345331350)  [8 启动任务 9](#_Toc345331351)  [9 停用任务 9](#_Toc345331352)  [10 修改任务信息 9](#_Toc345331353)  [11 查看任务的运行状态 11](#_Toc345331354)  [12 删除指定任务（慎用！！） 12](#_Toc345331355)  [13 配置任务告警 12](#_Toc345331356) 图标说明 ：新增任务 ：删除选择的任务  操作栏中的图标  ：编辑当前任务  ：跳转到新建任务步骤页面  ：跳转到周期管理页面  ：启用选择的任务  ： 停用选择的任务  ：弹出任务告警配置页面  任务状态栏图标：  ：任务状态正常状态  ：任务状态停用状态 任务查询 点击左侧树形菜单：普通管理-->任务管理，打开页面如下：    上面为查询条件框，下面为已有任务列表展示框，支持分页，按照任务的更新时间排序。其中表格内展示了任务基本字段，想查看具体任务信息，点击或者双击任务行，在弹出的页面中查看或者修改任务详细信息。   任务名模糊查询与精确查询 查询条件框的任务名称默认是模糊查询的，可以利用它搜索名字相仿的任务。    如果不想模糊查询，可以在任务名后面补充”;”号进行精确查询。   新建任务 点击新增任务按钮，打开页面如下：    其中字段带\*标识的为必填字段，有些字段给出的值是默认值，置灰的字段是由是否有多批标志字段决定的，各字段都给出了限制和校验，任务信息录入完毕，各字段符合要求，点击保存，会提示保存成功    同时任务列表框中第一行会展示刚新建好的任务信息。新建好的任务默认都是停用的。 任务的主要配置项 *任务优先级*：不建议修改，普通管理员无权修改。数值越小，优先级越高。TCC调度的最小单元是任务周期，任务周期的优先级继承自任务。所以优先级越高的任务，它的任务周期优先从等待队列调度到运行队列中运行。  *任务名称*：因为任务名支持模糊查询，如果TCC的分类不够用的话，可以利用名称来实现，比如使用相同的前缀或者后缀。  *任务执行周期类型和周期长度*：关键项，修改需要很谨慎。TCC配置的基准时间（默认配置2012-01-01）加上周期类型和周期长度就可以确定任务的每个周期的周期ID（周期Id是周期的起点时间点使用“yyyyMMdd-HH”格式化的字符串表示），比如周期类型为小时，周期长度为12，那么任务的周期Id序列就是20120101-00、20120101-12、20120102-00、20120102-12等等。目前支持4种周期类型，小时、天、月、年。任务启动后不允许修改，停止时可以修改，但是修改会删除原来运行产生的任务周期记录，重新初始化新的任务周期记录。  *周期偏移*：任务周期通常使用周期Id表示，而任务周期的可以运行时间点是周期Id对应的时间加上偏移时间。格式为xMxD xhxm，即偏移x月x天 x小时x分钟。例如周期Id为20120725-12，而它的任务的周期偏移是0M0D 3h30m，那么改任务周期的可以开始运行的时间点是2012-07-25 15:30:00。  *任务依赖关系*：为空表示忽略依赖关系，否则请按照如下格式“任务名,是否全周期依赖(0或1),是否忽略错误(0或1);...例如：task1,0,0;task2,0,1”填写，只能依赖具有可见权限的其它任务，否则权限不足。  任务的依赖关系本质是也是在任务周期中起作用的，如下：  大周期任务依赖小周期任务时，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。  *周期间是否顺序依赖*：任务的当前周期是否依赖任务的前一个周期，一般选择“是”，否则同一个任务的不同周期可能并行运行。  *是否有多批标志*：是否是需要处理多个文件的多批次任务（这里将一个文件看做一个批次），如果是非多批任务，则TCC默认创建一个批次，如果是多批任务，TCC会根据*分批结束标识*查询*文件列表*中配置的文件，选择属于当前周期的文件，并为每一个文件创建一个批次。所以如果是多批任务，则需要配置*分批结束标志、输入文件列表*。  *分批结束标志*：决定了TCC创建任务周期的批次的结束条件。如果是“普通方式(任务执行逻辑结束)”，直接当做非多批次任务处理；如果是“指定的输入文件处理处理完成”，需要配置*输入文件列表*，当任务周期执行时仅通过ls命令查询一次属于当前周期的文件，为每个文件创建一个批次，然后TCC处理完任务周期的所有批次后结束运行；如果是“等待时间内输入的文件处理完成”， 需要配置*输入文件列表*和*等待输入时间，*当任务周期执行时以固定时间间隔查询属于当前周期的文件，为每个文件创建一个批次直到超过了*等待输入时间*，然后TCC处理完任务周期的所有批次后结束运行；如果是“超过等待时间，且最少处理N个文件”， 需要配置*输入文件列表*、*等待输入时间*和*输入文件最少个数，*当任务周期执行时以固定时间间隔查询属于当前周期的文件，为每个文件创建一个批次直到超过了*等待输入时间*而且已经创建的批次数超过*了输入文件最少个数*，然后TCC处理完任务周期的所有批次后结束运行；如果是“超过等待时间，或者最少处理N个文件”， 需要配置*输入文件列表*、*等待输入时间*和*输入文件最少个数，*当任务周期执行时以固定时间间隔查询属于当前周期的文件，为每个文件创建一个批次直到超过了*等待输入时间*或者已经创建的批次数超过*了输入文件最少个数*，然后TCC处理完任务周期的所有批次后结束运行；  *输入文件列表：*文件的全路径，全路径中可以包含\*(任意批次)或者?(单字符匹配)通配符，亦可以使用$OS\_USER和$OS\_GROUP变量（TCC会自动替换成相应的OS用户和OS用户组），同时注意真正推送到服务器上的文件名中必需包含时间信息，即带上“***\_yyyyMMddHH***”*，例如：/home/hadoop/HiBI/ods/data/ODS\_UP\_USER\_INFO\_DM/up\_ODS\_UP\_USER\_INFO\*.txt。*  *部署节点：*多选，顺序很重要。每个任务可以配置多个部署节点（这些节点的任务部署信息完全相同，包括连接的用户名,rsa密钥,部署的任务脚本，需要处理的ods文件等等），TCC在节点执行相应脚本出错（连接主机或者认证出错直接退出，其它情况（如：脚本执行出错，执行中网络断连）按照步骤的重试次数重新运行多次）时会在下一次调度后切换到下一个节点执行，直到运行成功、超时或者、文件未到达退出结束；  *执行OS用户*：关键项。执行任务步骤命令所使用的操作系统用户，同时用于任务操作权限的控制，对应$OS\_USER变量值；  *所属用户组*：OS用户所属的操作系统用户组，不可编辑，同时用于任务可见权限的控制，对应$OS\_GROUP变量值；  *所属业务*：用户组隶属的业务，不可编辑；  *任务最早起始时间*：当配置好*任务执行周期类型和周期长度*后就可以确定任务的周期Id序列*，*会参与调度和执行的任务周期为原序列中处于*任务最早起始时间*之后的所有周期Id，之前的任务周期全部忽略，当考虑依赖关系时，处于*任务最早起始时间*之前的周期也忽略。  *重做类型*：仅对单批次、顺序依赖的每小时或者每天任务有效，包含集成重做、周末重做、月末重做三种类型；如果是“集成重做”，任务配置了*重做开始与结束时间*以及*集成重做合并的天数*N（非0），那么就将每N天内的历史任务周期（需要重新初始化）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而且调用脚本时传递的周期范围为N天内的第一个周期与当周内最后一个周期，其它的任务周期的均虚拟执行）；如果是“周末重做”，任务配置了*重做开始与结束时间*，那么就将每周内的历史任务周期（需要重新初始化）作为一个任务周期执行（TCC实际执行每周内的最后一个任务周期而且调用脚本时传递的周期范围为当周内第一个周期与当周内最后一个周期，其它的任务周期的均虚拟执行）；如果是“月末重做”，任务配置了*重做开始与结束时间*，那么就将每月内的历史任务周期（需要重新初始化）作为一个任务周期执行（TCC实际执行每月内的最后一个任务周期而且调用脚本时传递的周期范围为当月内第一个周期与当月内最后一个周期，其它的任务周期的均虚拟执行）；其它的重做相关配置，如当前周期处于*重做开始与结束时间之外，*或者*重做类型*为“集成重做“而且*集成重做合并的天数为*0，当前周期都会正常运行；  任务在重做时，利用集成重做、周末重做、月末重做，可以减少周期的步骤的执行频率，以帮助执行命令有机会处理更大跨度的周期范围，从而提高重做的效率。当然，要使用这三种重做方式，执行命令必需同时支持起始周期Id和结束周期Id这两个参数。  *任务周期占用的资源数*：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； 新建任务步骤页面 建好任务后，点击，跳转到新建步骤页面    其中表头显示了当前传入的任务ID，下排按钮分别为：新增、上移、置顶、下移、置底、删除，点击新增任务，弹出新增任务页面，任务ID自动传入，步骤Id后台生成，填入字段，如下：  步骤默认是启用的，点击保存按钮，给出保存成功提示，同时任务步骤列表相应变化    可以继续新增步骤，保存成功后如下    注意：TCC会按照步骤Id的顺序执行每个启用的步骤，所以对于有先后顺序约束的步骤，请确认好步骤Id的顺序，不符合要求可以点击上面的图标，进行相应的排序。此时新建任务信息（包含任务步骤）就完成了。回到任务列表页面就可以进行相应的操作。 任务步骤的主要配置项 *失败重试次数*：TCC执行步骤时，如果执行命令失败，会重新尝试执行的总次数；  *超时时间(分钟)：*TCC执行步骤时，如果执行的时长超过了设定的*超时时间*，那么步骤执行状态变为超时，批次和周期的状态也变为超时；  *执行命令：*TCC执行步骤时执行的命令，TCC会在远程登陆部署节点的OS用户后调用命令时，真正执行的命令格式为：  ***执行命令* [*批次文件名*] *起始周期Id 结束周期Id***  当任务为多批次任务时，[*批次文件名*]为每个批次对应的文件名；否则，[*批次文件名*]为空。  例如，userA用户配置了使用osuserA执行非多批任务A，任务A的步骤A的执行命令是/home/hadoop/HiBI/HIBI/exec-hive.sh dw\_up\_ul\_sms\_multidim\_stat\_dm.sql，TCC当前正在执行任务A的20120705-00周期，那么TCC执行周期的步骤A时，调用的linux命令是home/hadoop/HiBI/HIBI/exec-hive.sh dw\_up\_ul\_sms\_multidim\_stat\_dm.sql 20120705-00 20120705-00;  需要注意的是：  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、为了能让TCC将执行日志展示到页面上，脚本应该至少将运行日志信息输出到控制台。 启动任务 点击左侧树形菜单：普通管理-->任务管理，点击任务行的“操作”栏中按钮，此时会有两种情况：   1. 此任务并没有可用的任务步骤，此时会弹出警告框     选择“是”会跳到任务步骤列表页面，然后新增或者启用任务步骤，完成后再回到任务管理页重新点击任务行的按钮可启动任务；选择“否”会启用该任务，此时变成，任务状态由停用变成正常。   1. 有可用的任务步骤，会直接启用，此时操作栏中变成，任务状态由停用变成正常。   只有当任务启用后，任务的周期才会参与调度和执行。 停用任务 点击左侧树形菜单：普通管理-->任务管理，点击任务行的“操作”栏中按钮，任务状态由正常变成停用，同时后台会停掉任务的所有运行的任务周期。  停用任务后，TCC会停止掉等待队列或者运行队列中的任务周期，并停止掉任务周期的步骤关联的*命令进程。*  注意：TCC检测到任务停用需要一段时间，通常至少需要3秒，所以停用任务后，如果想要重新启用任务的话，请间隔较长时间再进行。 修改任务信息 点击左侧树形菜单：普通管理-->任务管理，先**停用任务**，然后再双击任务行或者点击任务行的按钮，弹出编辑任务页面，如下：    如果任务没有停用，会先弹出*周期类型*和*周期长度*不允许修改的提示框    否则直接弹出修改任务页面    修改完成后，点击保存，会提示保存成功，然后**启动任务**。要想修改任务步骤，点击跳到任务步骤按钮，点击修改按钮就可以修改任务步骤。  注意事项：1、修改*依赖关系*时，需要操作员具有所有依赖任务的可见权限；  2、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。**  3、如果不得已要修改周期类型或者周期长度，保存后将会删除原来运行产生的任务周期记录，重新初始化新的任务周期记录，如果不想重新初始化的历史任务周期运行，请记得调整*任务最早起始时间*。 查看任务的运行状态 点击左侧树形菜单：普通管理-->任务管理，点击任务行的按钮，会跳转到此任务的周期管理页面，默认显示该任务当月的所有任务周期：    详细请参见**周期管理**。 删除指定任务（慎用！！） 选择想要删除的任务行，点击，如果没有选择，提示至少选择一行记录，选择好后当前任务还在运行中，会提示你先停止该任务    当前任务允许删除，会弹出警告框，删除操作会删除任务、任务步骤、任务运行状态等和任务相关的信息，使用该操作请确定要删除的行数以及选择的任务名，谨慎！！！    点击确定，删除相应的任务，提示删除成功。 配置任务告警 点击左侧树形菜单：普通管理-->任务管理，点击任务行的按钮，会弹出告警配置页面，详细请参见**告警管理**。 |