|  |  |
| --- | --- |
| 产品名称 | 密级 |
| DataOne数据模型 | 机密 |
| 产品版本 | 共 页 |
|  |

工作总结

|  |  |  |  |
| --- | --- | --- | --- |
| 拟制 |  | 日期 | 2017/08/15 |
| 审核 |  | 日期 |  |
| 批准 |  | 日期 |  |



华为技术有限公司

版权所有 侵权必究

（仅供内部使用）

# 

[工作总结 1](#_Toc490591687)

[1. ID调研基础 3](#_Toc490591688)

[1.1 调研介绍 3](#_Toc490591689)

[1.2 调研步骤理解 3](#_Toc490591690)

[1.2.1 调研需求确认 3](#_Toc490591691)

[1.3 调研工作展开 4](#_Toc490591692)

[2. 模型设计基础 4](#_Toc490591693)

[2.1 模型设计介绍 4](#_Toc490591694)

[2.2设计步骤理解 5](#_Toc490591695)

[2.3设计工作展开 5](#_Toc490591696)

[3. 相关工作说明 6](#_Toc490591697)

[4. 业务方面熟悉 6](#_Toc490591698)

[5．各类权限申请审批流程总结 6](#_Toc490591699)

[5.1 SVN权限申请 6](#_Toc490591700)

[5.1.1 权限申请地址 6](#_Toc490591701)

[5.1.2 选择“TC-Cloud-DataServiceTM” 7](#_Toc490591702)

[5.1.3 选择“申请加入群组” 7](#_Toc490591703)

[5.1.4 填写资料 7](#_Toc490591704)

[5.1.5 SVN安装 8](#_Toc490591705)

[5.2 VPN权限申请 9](#_Toc490591706)

[5.2.1 网络安全考试 9](#_Toc490591707)

[5.2.2 权限申请地址 9](#_Toc490591708)

[5.2.3 申请使用步骤 10](#_Toc490591709)

[5.3 华为UMA申请 12](#_Toc490591710)

[5.4 代运维权限 12](#_Toc490591711)

[5.5 Proxy代理权限 13](#_Toc490591712)

[5.6 元数据管理系统访问方法 14](#_Toc490591713)

# ID调研基础

## 调研介绍

Id调研是针对ODS进行工作，主要内容包括系统级调研、表级调研和字段级调研，目前工作主要集中在表级调研和字段级调研。通过表级调研，明确ODS表名及中文名、所属业务模块、加载算法、推送时间等；字段级调研，明确ODS表各个字段名及中文名、源数据类型、是否主键、是否个人数据、加密函数、代码值枚举等，而调研最主要的目的就是为了使下一阶段的模型设计工作流程能够稳定准确的进行。

## 调研步骤理解

调研就是为了搞清楚需要设计的表的结构信息、业务场景、数据量以及表与表之间的关联关系等，为ODS到DWD的结构设计做准备的过程。其中主要牵涉到以下几个文档的更新维护：《02tableLvlTask.xls》、《03fieldTask.xlsx》、《personal\_privacy\_data\_summary.xlsx》

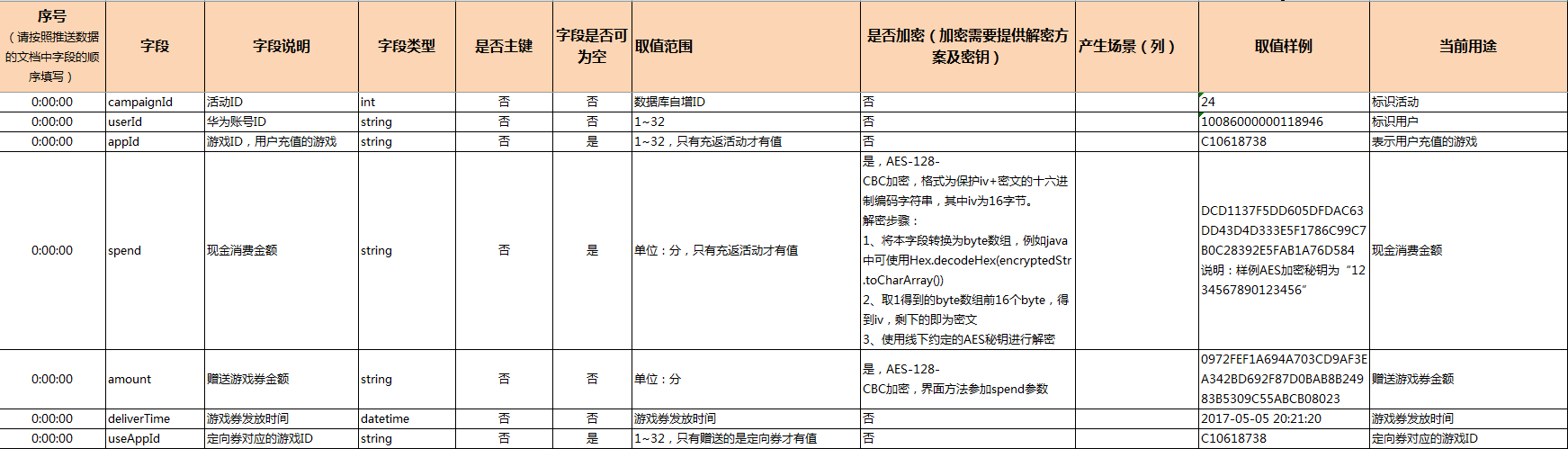
（1）《02tableLvlTask.xls》是表级调研文档，统计表的包括名称、业务模块以及所DWD所属主题等；

（2）《03fieldTask.xlsx》是字段级调研文档，统计表的包括字段名称、字段的类型、业务场景、主外键、关联关系等；

（3）《personal\_privacy\_data\_summary.xlsx》个人信息文档，统计表中是否存在个人隐私数据需要进行加密处理，如：手机号、imei、ip等。

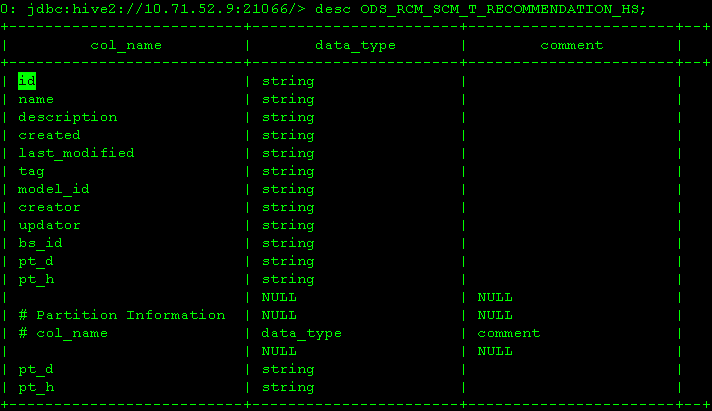
### 1.2.1 调研需求确认

调研需要找是否有TALM单，拿到TALM单号后进行查询需求内容（http://talm.huawei.com/ipd/t\_alm/#!alm/role/rolereport/roleReport.html?tab=1078），在TALM单中会附录数据源表的数据字典文档，当然如果数据接入后会存放在SVN中的“数据源接入”文件夹中，从数据字典文档中能够清晰的看到ODS源表的数据结构（如果没有数据字典，请找PM或者需求提出人获取）。



## 调研工作展开

（1）使用跳板机进入ODS源表环境（C30、C60）查看该表是否已接入源数据；



（2）从根据表级文档和字段级文档的内容来进行相关调研，最主要是要搞清楚需要调研的表的使用场景、主外键、与其他表的关联关系、取值范围，便于后期进行DWD层相关表的设计、开发。





# 2. 模型设计基础

## 2.1 模型设计介绍

模型设计是针对调研结果的基础上进行建模，将ODS的数据整合以供后期汇总应用。而模型设计最常见的是将调研结果设计到整合层中（模型总体架构：贴源层 -> 整合层 -> 汇总层 -> 应用层），设计结果包括：表名、字段名、字段类型、加解密处理、数据清洗、大字段拆分等。比如日期要统一数据格式；个人数据作统一加密处理；字段命名统一并沿用历史命名；大字段拆分出业务需要的字段等。

## 2.2设计步骤理解

设计需要比较详细的调研结果进行支撑，如果调研结果不详细或者数据源没有接入会对设计结果产生误差，可能造成命名误差、业务场景误差、主题域设计误差等，设计过程中主要更新维护的文档：《PDM.xlsx》、《词根.xlsx》、《通用代码表.xlsx》

（1）《PDM.xlsx》是已入仓ODS表对应的整合层设计结果，新设计的表都需要维护到该文档；

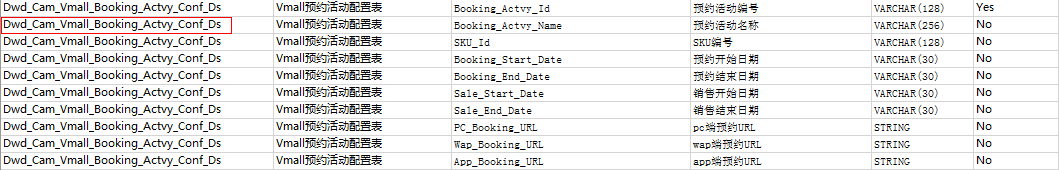
（2）《词根.xlsx》所有在表设计阶段是用的字段都是由词根组成的，没有能够表达字段意义的词根需要新增词根；

（3）《通用代码表.xlsx》在设计表的同时，有一种特殊的字段，能够被穷举并赋予对应枚举值确切含义的集合就是代码值，这种字段需要维护进通用代码表。

## 2.3设计工作展开

（1）首先需要根据调研结果对新设计的表进行命名，而在命名之前还需要将表进行主题域的划分，主题域包括：参与者(Pty)、产品(Pro)、协议(Agt)、设备(Eqp)、

事件(Evt)、位置(Loc)、营销(Cam)、内容(Con)、财务(Fin)、销售(Sal)、ONLINE、参数(Ref)、代码(Cde)，这13大主题域，而表的命名规范为：DWD\_域名\_表名称\_结尾（结尾分为hm,dm,ds,ds\_his）。一般来说订单归属Sal，广告归属Cam，，日志类一般归于Evt，用户的基础信息归于Pty，手机终端相关的产品信息归于Eqp。主题后面的二级命名一般沿用业务习惯命名，比如应用市场hispace，广告adv，Vmall用vmall等，二级命名后面一般使用源表名，能够描述清楚表的意思或用途即可，最后的结尾习惯\_ds或\_dm，dm一般指日期增量的表，ds一般指全量表。表的全部字长不建议超过40位，如果有超过可适当缩减。



（2）字段命名原则上是遵循词根进行翻译的，同时字段命名尽量考虑和已入仓的整合层pdm表及字段的关系，同样内容的字段或者同一主题下的表的字段命名尽量一致，简单的来说就是先到Pdm里找类似的表有没有这样的字段名称，有直接用字段名称，中文注释，字段类型，没有要么erwin里定义进去系统自动生成，要么自己根据词根来定义个，大致来说遵循：'编号类：以“编号”(\_ID)结尾，指数据项被赋予具有一定规律、易于计算机和人识别处理的符号，形成编号元素，每个编号均对应特定的业务对象，适用于具有特定编号规则的信息项，例如消费者编号、产品编号、设备编号等；代码类：以“代码”(\_Cd)结尾，指数据项以对各取值范围进行结构化处理的方式体现，每个代码取值只代表业务对象的一种可能性，适用于有明确分类、分组或多取值的信息项，例如行业代码、语言代码、性别代码、消费者类型代码等； 指示器类：以“标志”(\_Flg) 结尾，指数据项仅包括“是”、“否”两种结果，1 表示是，0表示否，例如：外部产品标志，花粉用户标志，会员标志等；文本类：指数据项以纯文本的形式体现，适用于很难结构化的说明描述类信息，例如客户名称、产品描述等；一般以“描述”，“名称”，“说明”等结尾金额类：一般以“金额”(\_AMT)，“余额”(\_BAL)结尾，指数据项以货币金额的形式体现，适用于各类财务信息，例如违约金额、账户余额等；

（3）代码字段除特殊情况都要求是VARCHAR(16)，同时要求必须先到通用代码表中通过字段命名和码值内容确认有没有现成的代码内容，有直接使用当前的代码字段命名，没有则需要同步在通用代码表.xls中增加。

# 3. 相关工作说明

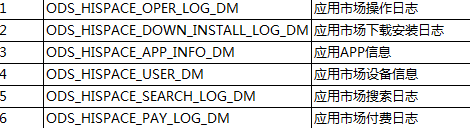
国内数据源参与调研和设计50余张，目前每周前三天进行新需求的调研和设计工作，后两天完善香港数据源cloud\_folder十余张的调研和设计；输出了元数据系统相关登录方法以及各类权限的申请方式文档（见第5节）；目前正在着手开发方面的知识的补充。

# 4. 业务方面总结

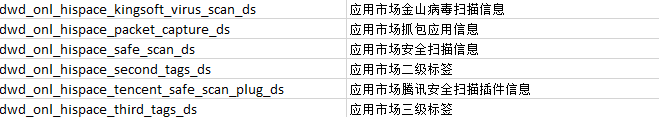
主要负责熟悉的业务是应用市场和游戏中心，总结了两块业务的已入仓的ODS和DWD的对应关系以及市场场景。

应用市场方面：

（1）常用的一些表

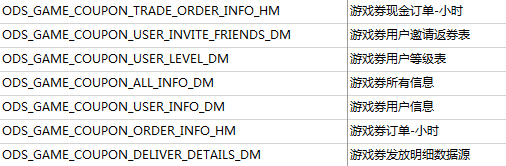


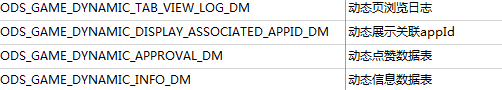
（2）应用市场的ODS囊括了用户的操作日志、搜索日志、下载日志、曝光日志、搜索日志、付费日志、点赞评论、评论回复、用户评分、设备编码参数等表，旨在用户可以在应用市场上搜索、下载、管理、分享最贴心的移动应用，安全机制方面首创“开发者实名认证+3重系统检测+人工专员复检”，其相关的一些表为

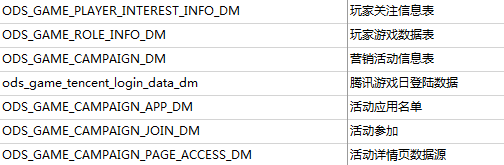


游戏中心方面：

（1）游戏中心目前入仓的业务主要集中在游戏券信息、活动、订单、游戏玩家数据、游戏营销活动、游戏日登录数据等；而游戏中心提供最新游戏下载、活动礼包、正版游戏推荐等功能。







（3）应用市场以及游戏中心的dwd分布



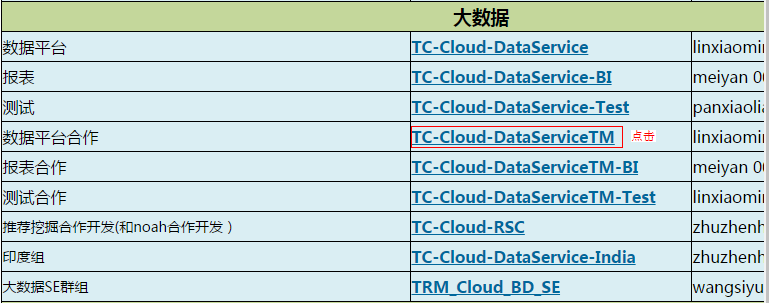
# 5．各类权限申请审批流程总结

## 5.1 SVN权限申请

### 5.1.1 权限申请地址

<http://3ms.huawei.com/hi/blog/690361_1962991.html>

### 5.1.2 选择“TC-Cloud-DataServiceTM”



### 5.1.3 选择“申请加入群组”



### 5.1.4 填写资料

“资产/群组所在部门/产品”选择”资源线”并默认部门，然后填写相关资料并提交申请，等待linxiaoming 00216915审批



### 5.1.5 SVN安装

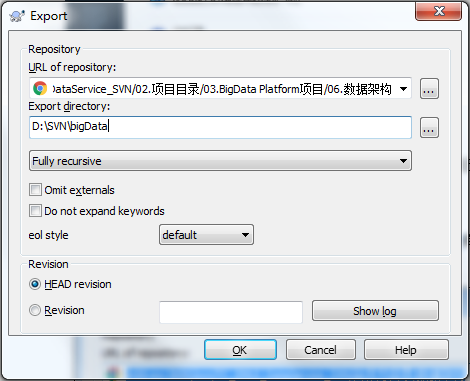
（1）权限审批通过后，在iDesk的软件栏搜索SVN，然后下载安装。（64位的操作系统安装64位，32位的操作系统安装32位，去系统属性看是32位还是64位）



（2）新建一个文件夹，例如：bigData，然后右击文件夹bigData，再点击SVN checkout，

出现以下图，在URL of repository填：

https://szxsvn06-ex:3690/svn/TC\_EMUI\_DataService\_SVN/02.项目目录/03.BigData Platform项目/06.数据架构



再点击OK。

## 5.2 VPN权限申请

### 5.2.1 网络安全考试

1.申请之前要通过【消费者BG网络安全服务上岗证考试】

（1）参加考试链接：

<http://ilearning.huawei.com/exam/paper.builduserpapercurr.flow?sid=4805090685208887>

（2）查询考试分数

<http://ilearning.huawei.com/exam/userpaper.viewuserhispaperlist.flow?p_id=4805090685208887>

（3）证书下载链接 （电子证书会在考试通过后2-3天生成，并邮件通知下载【用UE浏览器】）

<http://ilearning.huawei.com/ilearningportal/ilearningPortal.html#!ilearningPortal/myiLearning/myiLearning.html?code=1>

### 5.2.2 权限申请地址

<https://opene2.hicloud.com/portal/openservice/index/homepage>

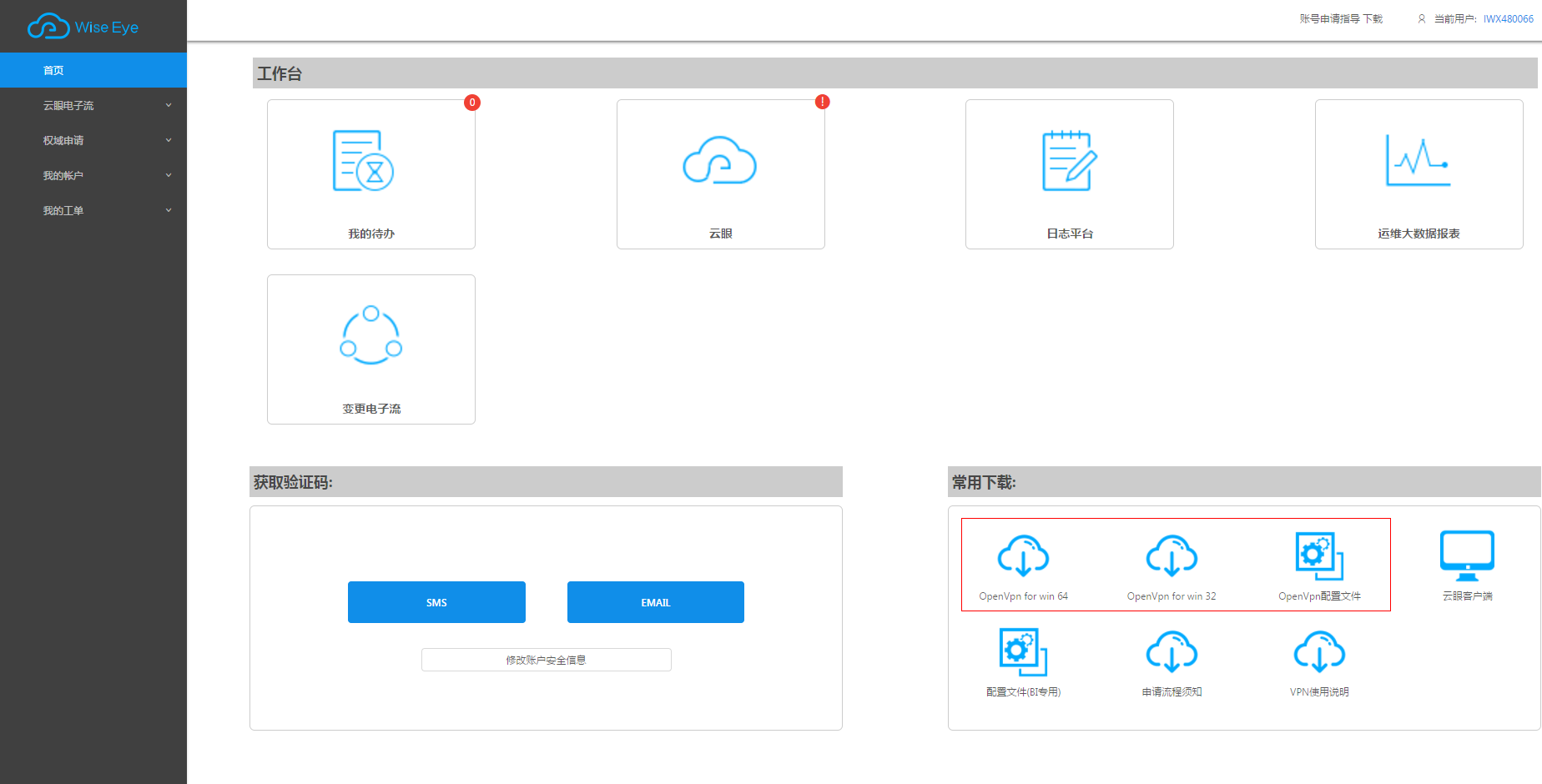
### 5.2.3 申请使用步骤

（1）选择“申请”



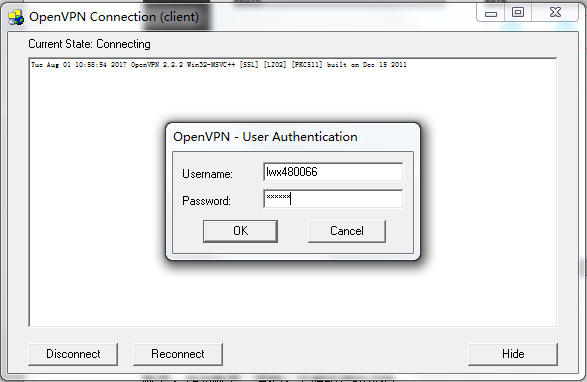
（2）安装vpn

vpn客户端、配置文件此处下载

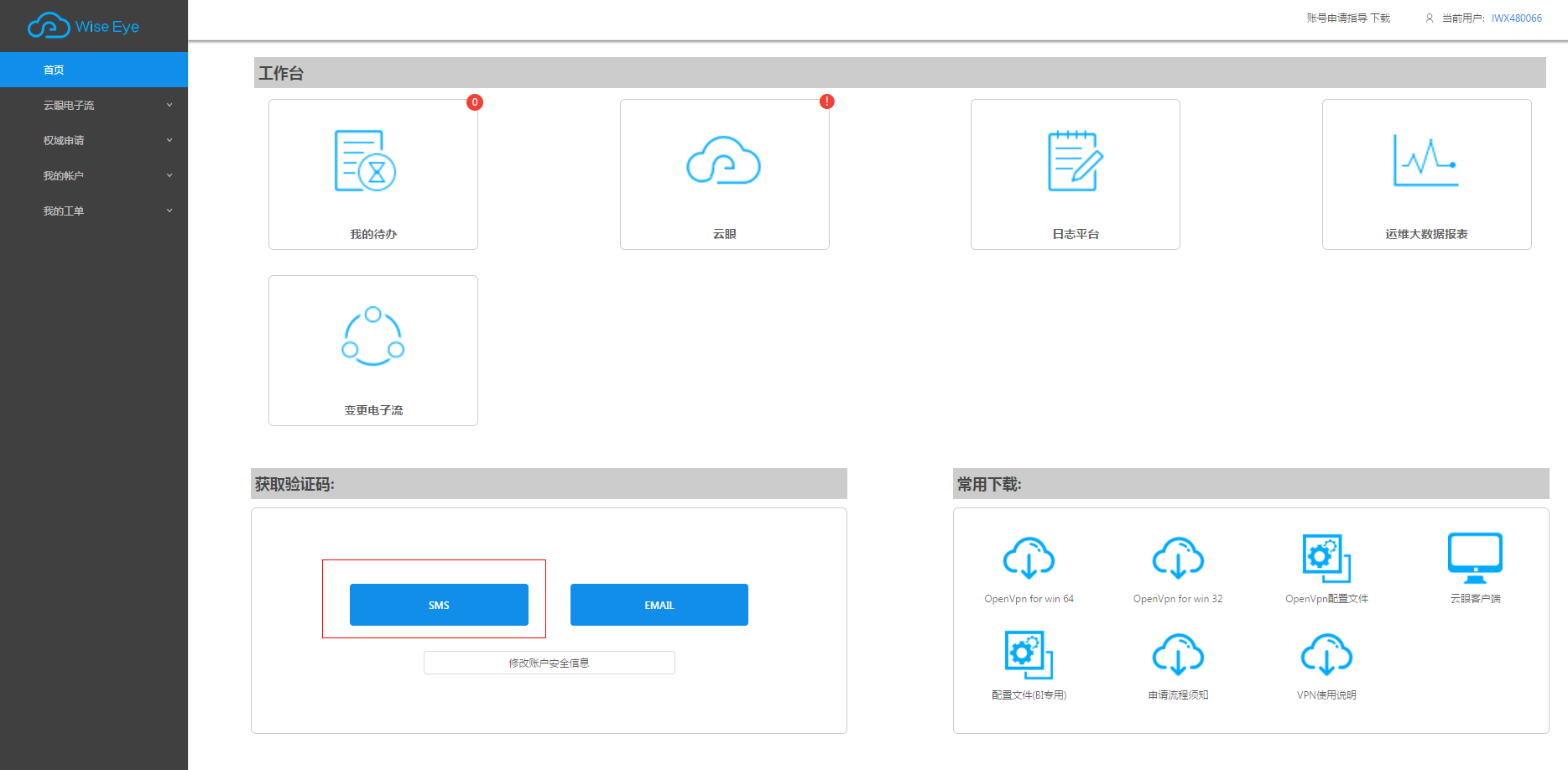


（3）登录VPN

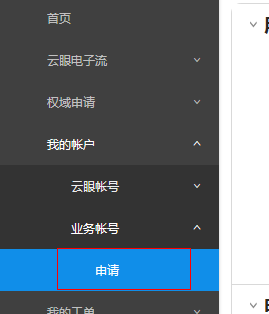
账号：华为账号 密码：云眼短信验证码

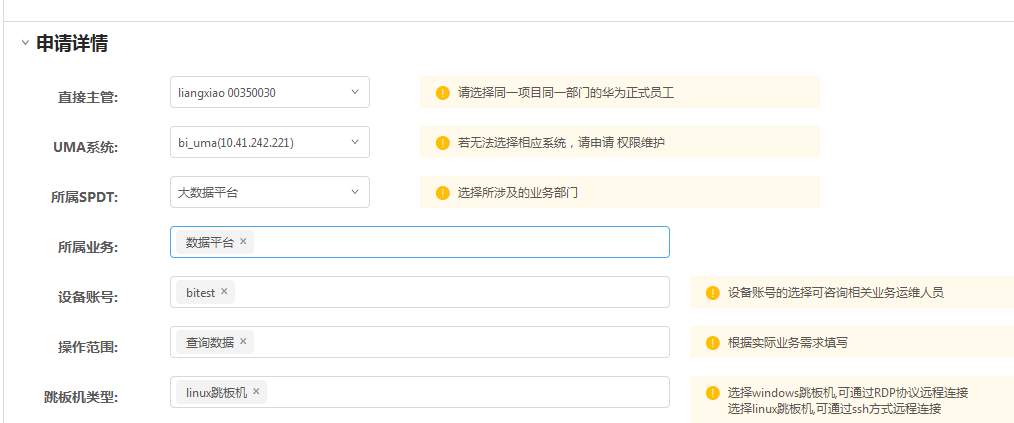


注：云眼短信密码点击此处发送：



## 5.3 华为UMA申请





## 5.4 代运维权限

iAccess右键🡪Internet访问阻断日志：



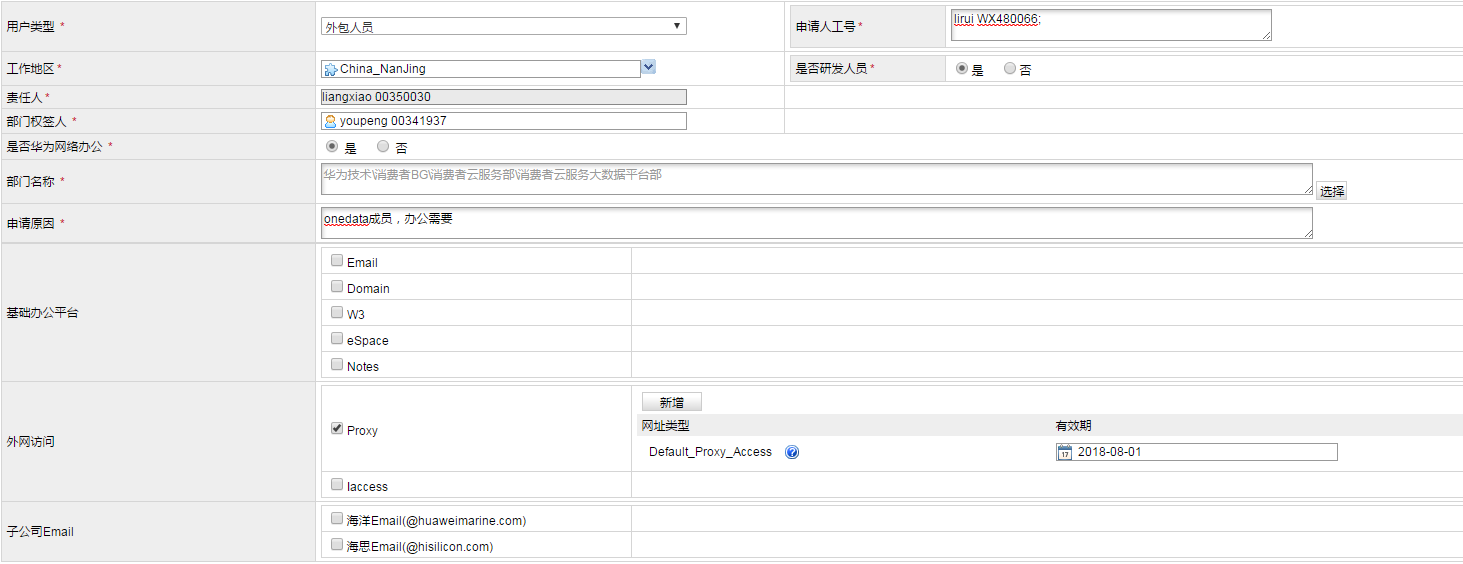


## 5.5 Proxy代理权限

注：只需申请default权限

申请地址：

<http://w3.huawei.com/eso/hidm/basicApp/basicOffice.do?method=basicOffice&catalogId=10137>



## 5.6 元数据管理系统访问方法

（1）通过BIPortal链接登录：<https://portal.data.hicloud.com/BIPortal/>

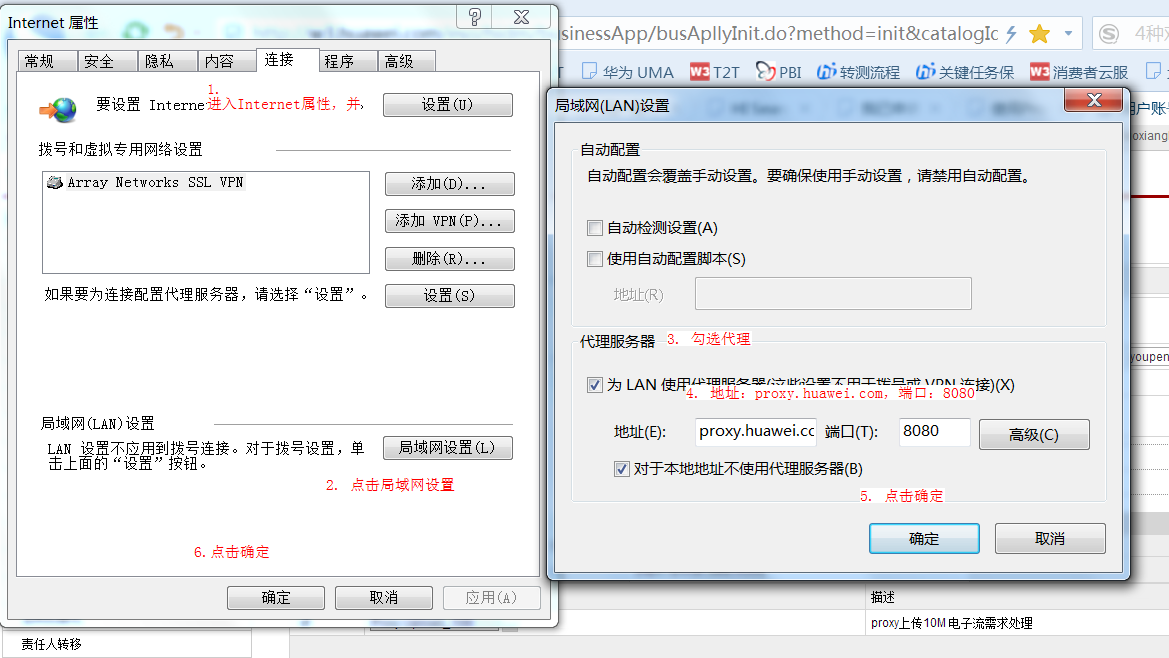


（2）通过OMS链接登录：<https://portal.data.hicloud.com/oms/>



（3）接入步骤：连接iAccess 🡪 选择”互联网关闭” 🡪 IE中勾选”代理服务器” 🡪 访问元数据管理系统

注：访问元数据管理系统需要外网权限，既需要在IE中打开代理服务



（4）元数据管理系统访问方法及权限申请文档总结

