**终端设备历史开发手册**

1. **数据字典**

**1.1表级**

英文名：dwd\_eqp\_device\_ds\_his

中文名：终端设备历史表

表类型：拉链表

ETL调度频率：日表

**1.2字段级**

|  |  |  |  |
| --- | --- | --- | --- |
| 字段名 | 字段类型 | 字段中文名 | 备注 |
| imei | VARCHAR(128) | 设备编号 |  |
| start\_date | VARCHAR(8) | 开始日期 |  |
| hw\_device\_flg | SMALLINT | 华为设备标志 |  |
| hw\_device\_type | VARCHAR(256) | 华为设备类型 |  |
| device\_desc | VARCHAR(256) | 设备描述 |  |
| bom\_id | VARCHAR(128) | BOM编号 |  |
| ship\_box\_id | VARCHAR(128) | 发货箱单编号 |  |
| did | VARCHAR(128) | 设备唯一号 |  |
| inside\_name | VARCHAR(128) | 内部型号 |  |
| device\_name | VARCHAR(128) | 外部型号 |  |
| prod\_name | VARCHAR(256) | 产品名称 |  |
| pdt | VARCHAR(128) | PDT |  |
| brand | VARCHAR(256) | 品牌 |  |
| tcsm\_first\_report\_time | VARCHAR(30) | 进销存首次上报时间 |  |
| tcsm\_first\_start\_region\_cd | VARCHAR(16) | 进销存首次开机行政区划代码 |  |
| series\_name | VARCHAR(256) | 系列名称 |  |
| external\_series\_name | VARCHAR(256) | 外部系列名称 |  |
| hota\_device\_flg | SMALLINT | HOTA设备标志 |  |
| preinstall\_emui\_ver | VARCHAR(128) | 预装EMUI版本 |  |
| preinstall\_rom\_ver | VARCHAR(128) | 预装ROM版本 |  |
| currt\_emui\_ver | VARCHAR(128) | 当前EMUI版本 |  |
| currt\_rom\_ver | VARCHAR(128) | 当前ROM版本 |  |
| device\_active\_time | VARCHAR(30) | 设备激活时间 |  |
| plmn | VARCHAR(256) | PLMN |  |
| etl\_time | VARCHAR(30) | ETL时间 |  |
| honor\_flg | SMALLINT | 荣耀标志 |  |

1. **数据来源**

**2.1．来源于17个业务系统:**

adv(广告) :dwd\_cam\_adv\_req\_log\_dm(广告请求日志)

bdreporter(rom采集) :dwd\_evt\_bdreporter\_device\_static\_info\_dm(BDREPORTER设备静态信息)

music(华为音乐) :dwd\_evt\_hw\_music\_interface\_api\_log\_dm(华为音乐接口调用日志)

movie(华为视频) :dwd\_evt\_hwmovie\_user\_access\_log\_dm(华为用户访问日志)

memberright(会员权益) :dwd\_pty\_hw\_membr\_rights\_ds\_his(华为会员权益)

hitop(主题) :dwd\_evt\_theme\_client\_register\_info\_dm(主题客户端注册信息)

hispace(应用市场) :dwd\_evt\_hispace\_device\_dm(应用市场设备),dwd\_evt\_hispace\_online\_game\_device\_dm(应用市场网游设备)

hicloid(基础云) :dwd\_evt\_cloud\_serv\_oper\_dm(云服务操作日志)

cfolder(云文件夹) :dwd\_evt\_cfolder\_user\_register\_dm(云文件夹用户注册)

bisdk(大数据采集) :dwd\_evt\_bisdk\_visit\_dm(BISDK访问用户)

youkuvideo(优酷视频) :dwd\_evt\_hwmovie\_interface\_api\_log\_dm(华为视频接口调用日志)

hota(设备升级) :dwd\_evt\_hota\_device\_chk\_log\_dm(HOTA设备检查日志)

wlan(天际通) :dwd\_eqp\_wlan\_open\_user\_ds(WLAN开通用户)

up(账号) :dwd\_up\_user\_device\_info\_ds(华为账号与IMEI绑定关系表),dwd\_evt\_up\_oper\_log\_dm(华为帐号操作日志)

trade(支付) :dwd\_sal\_order\_pay\_ds(订单支付表)

phoneservice(手机服务) :dwd\_evt\_phoneservice\_qstn\_receivesurvey\_log\_dm(问卷调查用户答案提交日志)

push(推送) :dwd\_eqp\_push\_user\_rout\_lnk\_dm(PUSH用户路由连接),dwd\_evt\_push\_log\_sign\_out\_log\_dm(PUSH登陆登出日志)

hiboard(负一屏) :dwd\_evt\_hiboard\_qry\_log\_dm(负一屏查询日志视图)

**2.2．其他依赖的表:**

dwd\_eqp\_dev\_id\_mapping\_ds(下文简称：mapping表):生产发货数据，用于获取bom\_id，did，设备描述字段，并判别是否华为设备

dwd\_eqp\_hw\_device\_bom\_ds(下文简称：bom表):bom数据，用于获取外部设备型号，内部设备型号，系列名，产品名等属性字段

dwd\_evt\_bisdk\_customize\_dm(下文简称：gps信息):GPS上报天气信息，用于获取行政区划字段

**2.3．备注:**

dwd\_evt\_up\_oper\_log\_dm，dwd\_evt\_push\_log\_sign\_out\_log\_dm两个源由于ODS推送时间过晚，故延迟一天入库。

1. **数据逻辑**

**3.1．主体逻辑**

通过17个业务源系统上报的imei信息，汇总华为业务所有设备信息。并且通过生产发货数据讲同一个设备的imei串联起来(did相同的imei表示是同一个设备)

**3.2．字段逻辑**

| 字段名 | 处理逻辑 |
| --- | --- |
| imei | 主键。来源于17个源系统。格式为：14到16位的数字和小写字母。参见本文4.1，4.5章节 |
| start\_date | 拉链表逻辑字段，标识整条记录状态的开始时间 |
| hw\_device\_flg | 关联生产发货数据(mapping表)，当此imei存在于mapping表时则认定为华为设备，否则为非华为设备  1，华为设备 0，非华为设备 |
| hw\_device\_type | 标识华为设备的类型，如手机，平板等  来源于bom表，(详见bom表开发手册)。部分用法参见本文4.4章节 |
| device\_desc | 描述型字段，来源于mapping表 |
| bom\_id | Bom编号，用于关联获取bom信息的外键。来源于mapping表 |
| ship\_box\_id | 发货箱单号，用于关联获取发货信息的外键。来源于mapping表 |
| did | 设备唯一标志。用于串联同一设备的不同imei（当某些imei的did相同时表示这些imei为同一设备）。来源于mapping表 |
| inside\_name | 内部型号。来源于bom表(详见bom表开发手册) |
| device\_name | 外部型号。首先取生产发货的外部型号，如果获取不到则通过17个源上报的信息来获取外部型号。具体获取优先级如下：bom>hota>push>hitop>phoneservice>wlan>music>bisdk>up>cfloder>adv>memverright>youkuvideo>bdreporter>hispace>hiboard。参见本文4.6章节。 |
| prod\_name | 产品分类。来源于bom表(详见bom表开发手册) |
| pdt | Product Development Team，产品开发团队。来源于bom表(详见bom表开发手册) |
| brand | 品牌分类。来源于bom表。当hw\_device\_flg标识为华为设备时，通过关联bom表获取品牌，获取不到则标志为“其他”。否则用外部型号关联手工维表dwd\_ref\_device\_brand\_rela\_par\_ds获取此外部型号所对应的产品名称(此表中不包含华为的产品)。否则根据外部型号来判断非华为的品牌，如：外部型号 LIKE '%SAMSUNG%' 则品牌为 '三星'，具体规则详见脚本。 |
| tcsm\_first\_report\_time | 进销存首次上报时间。来源于hota和gps。主要供给进销存系统使用。取hota检查到和gps上报的最早的时间 |
| tcsm\_first\_start\_region\_cd | 进销存首次上报时间。来源于hota和gps。主要供给进销存系统使用。优先取hota检测到的最早的行政区划(通过解析ip地址获取)，否则取gps最早上报的时间。当某一设备首次上报7天后则冻结不在更新(不管此字段当前取的是hota还是gps还是NULL值)。参见本文4.2章节 |
| series\_name | 系列号。来源bom表 |
| external\_series\_name | 外部系列号。来源bom表 |
| hota\_device\_flg | 表识hota上报设备的类型。  0，非hota上报数据 1，手机和平板 2，数据卡 3，盒子 4，路由器  取自hota升级日志的product字段。并且不来自SP\_ARD\_APP\_COMMON局点。 |
| preinstall\_emui\_ver | 首次上报的emui版本号。取业务上报的最早的emui版本。业务优先顺序:hota>bdreporter>phoneserver |
| preinstall\_rom\_ver | 首次上报的rom版本号。取业务上报的最早的rom版本。业务优先顺序:hota>bdreporter>phoneserver>bisdk |
| currt\_emui\_ver | 当前的emui版本号。取业务最后上报的emui版本号。业务优先顺序同上。  标准化：数字.数字 如：5.1 ( 正则表达式：(\d+\.\d+) ) |
| currt\_rom\_ver | 当前的rom版本号。取业务最后上报的rom版本号。业务优先顺序同上 |
| device\_active\_time | 首次激活时间。取各个业务首次获取的最小时间。 |
| plmn | MCC+MNC组合。移动国家码和移动网络号。取自各个业务系统最后一次上报的信息。业务优先顺序:hota>bdreporter>bisdk>hispace>up>cfolder>push>hitop。  MCC码值请参考dwd\_cde\_common\_code\_ds表(cd\_en\_name = ‘MCC’)  MNC码值请参考dwd\_cde\_common\_code\_ds表(cd\_en\_name = ‘MNC’) |
| etl\_time | ETL时间字段。记录数据入库的服务器时间。 |
| honor\_flg | 是否荣耀设备。取自bom表。  1，是 0，否 |
| end\_date | 拉链表逻辑字段同时也是分区字段，标识整条记录状态的结束时间。当此字段为99991231时存储的是当前最新的全量设备信息，当为某一具体日期(如：20170501)时存储的是当天存在闭链(当天数据有更新)的数据。参见本文4.3章节 |

**3.3．ETL逻辑**

参考拉链表处理逻辑。

先在dwd\_eqp\_device\_ds\_src表中累积每个业务每天的全量数据，分区字段pt\_d:天分区 pt\_service:业务分区。然后取出终端设备历史表的前一天数据(start\_date<=’$last\_date’ AND end\_date > ‘$last\_date’)与dwd\_eqp\_device\_ds\_src当天全量数据(pt\_d = ‘$date’)做比较(主键关联)，若有字段有变动。则将前一天的数据的end\_date修改为$date，并且新插入一条记录(start\_date=$date, end\_date为99991231)。若当天的数据有昨天的数据没有则属于新增记录，在终端设备历史表中新增一条记录(start\_date为$date, end\_date为99991231)。

1. **其他**

**4.1．设备表中imei分类**

1，meid：14位 电信才有 一般为a开头：axxxxx 数字字母组合

2，imei ：15位 大多数为：86xxxxx，35xxxxxx 部分004xxxxx

3，SN ：16位 数据字母

其他的诸如mac地址，苹果的UUID等均不在终端设备表中

**4.2．进销存首次上报行政区划逻辑**

1，第一次入库若GPS有值，则直接冻结

2，第一次入库若Hota有值，则后面只有GPS的才更新，且只更新一次GPS即冻结

3，首次入库后7天直接冻结（不管是GPS或者hota或者直接为NULL）

4，GPS与GPS之间，hota与hota之间 不更新

5，“7天冻结”这个逻辑只判断GPS和HOTA分区，其他分区不参与

**4.3．拉链表几种简单的取数用法**

**4.3.1 取某天数据**

WHERE start\_date <=’$date’ AND end\_date>’$date’

**4.3.2 取当前最新的数据**

WHERE end\_date = ’99991231‘

**4.3.3取某天新增的记录(主键新增)**

**4.3.3.1方法1：**

正常方法取当前日期和前一天日期，两天的数据 做对比

**4.3.3.2方法2：**

SELECT

imei

FROM dwd\_eqp\_device\_ds\_his

GROUP BY imei

HAVING MIN(start\_date) = '$date';

**4.3.3.3 方法3：**

SELECT

imei

FROM dwd\_eqp\_device\_ds\_his t1

WHERE end\_date > '$date' AND start\_date = '$date'

AND NOT EXISTS (

SELECT 1 FROM dwd\_eqp\_device\_ds\_his t2

WHERE end\_date = '$date' AND t1.imei = t2.imei

);

**4.4．判断设备是手机还是平板**

手机：hw\_device\_type like '%手机%' AND hw\_device\_type NOT LIKE '%手机类备件%'

平板：hw\_device\_type like '%平板%'

**4.5． imei处理过程**

4.5.1 解密

4.5.2 标准化：revertDeviceId(deviceIdFormat(imei))

4.5.3 过滤非法imei判断是否非法imei : IsDeviceIdLegal(imei) AND LENGTH(imei)<17

**4.6 外部型号说明**

外部型号首先取自生产发货上报的数据，这部分数据是最准确的。如果取不到则取各个业务最早上报的外部型号。但是业务系统上报的数据质量参差不齐，所以有可能出现一些脏数据，这些属于正常现象，数据量不会太大。另外初始化数据也会使外部设备出现部分脏数据，目前采取的修复方式是取初始化以后业务首次上报的外部型号，如果初始化以后没有业务上报数据则依然取初始化数据，如果这个设备是活跃设备，外部型号总归会被上报数据刷新。