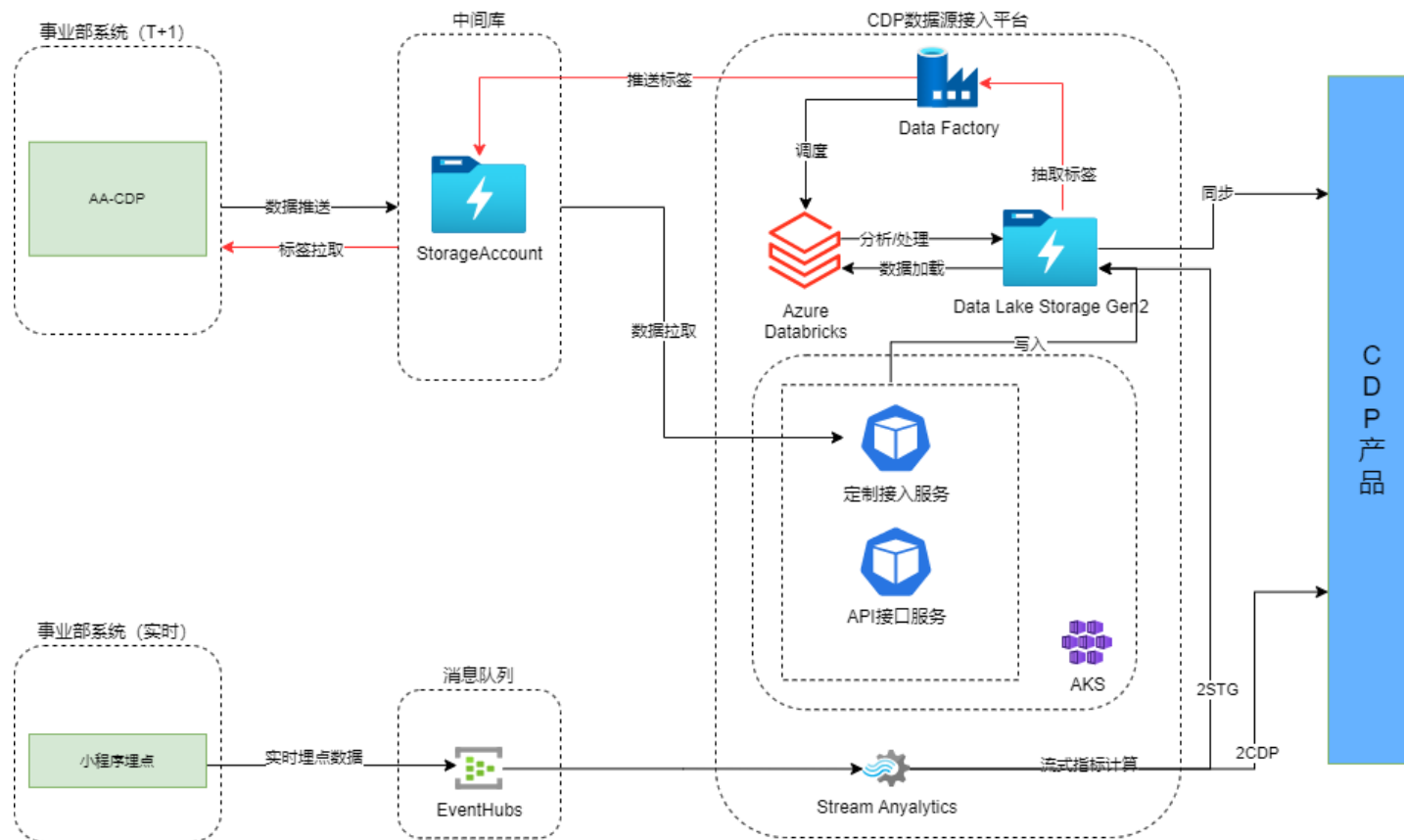


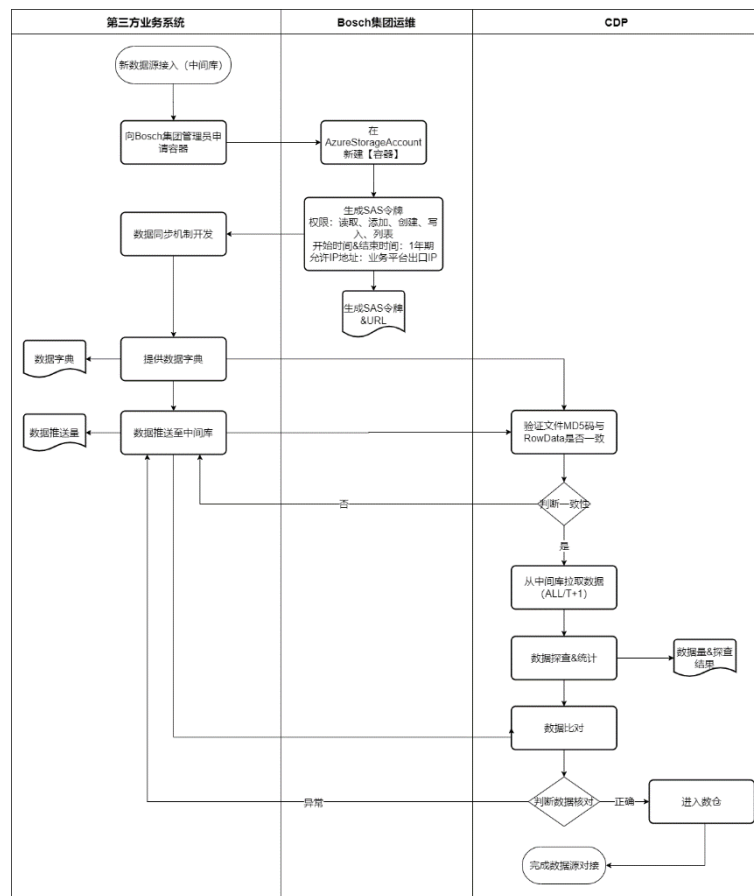
CDP – 事业部数据对接方案讨论

数据接入 设计思路



数据接入

中间库场景SOP及设计概要



- **隔离机制:** 为中间库创建独立的存储帐号, 每个数据源在此存储帐号中通过不同Azure Blob容器承载
- **存储机制:** 每个接口/表, 在当前容器根目录进行创建, 例如:
【/activateCustomerListSync】
- **分区机制:** 数据目录按照日期进行存放, 例如:
【/activateCustomerListSync/date=20161026】
- **数据文件结构:** Raw Data以csv格式保存, 分隔符采用竖线【|】分隔, 文件首行为字段Field (数据需要严格按照字段分隔); 数据文件已UTF-8格式进行编码;
- **认证与授权:** CDP平台配令牌及IP白名单
 - 权限【读取、添加、创建、写入、列表】
 - 开始时间&结束时间
 - 允许IP地址【业务平台出口IP】
 - SAS令牌举例
【<https://stcdpstorage.blob.core.chinacloudapi.cn/cmdtestdir?sp=racwl&st=2023-03-20T09:55:25Z&se=2023-03-20T17:55:25Z&spr=https&sv=2021-12-02&sr=c&sig=mPCaNHVhYU20QLgtLrS9riZLnSTRWp224fO7o1YZlvY%3D>】
- **一致性机制:** 第三方业务系统在上传RawData数据同步, 请同步上传对应的MD5校验码, 例如
【activateCustomerListSyncByActiveTime_20160831.md5】
- **异常处理机制:** 当数据比对发现异常时, 需联系第三方业务系统进行重新核对以及数据补推
- **数据字典规范:** 数据字典中需包含1) 表/接口名称; 2) 表/接口描述 (业务含义); 3) 字段名称; 4) 字段类型; 5) 字段描述; 6) 字段取值范围 (字典值);

数据接入

实时埋点数据

- 对接方式：小程序埋点

- CDP提供基于Azure Eventhub的endpoint

- 传输协议，方法：HTTPS (POST)
 - Header：Authorization: <SAS Token> (有效期1-2年)
 - 传输内容格式：待进一步讨论，JSON payload 比如：openid、unionid、触发事件、停留时长、路径等