利用“同态”加密
实现安全的数据交付

赵志刚

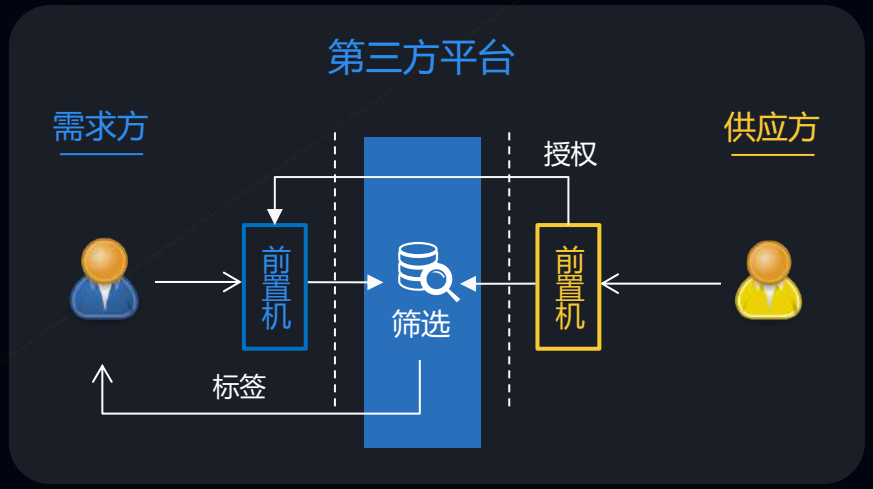
TalkingData 资深架构师

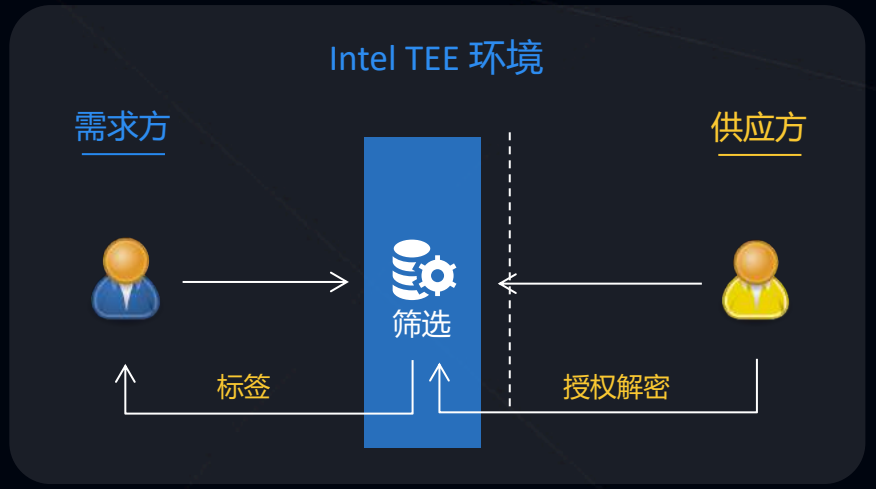
背景介绍

业内已有方案

信任第三方

- 第三方平台
- 第三方硬件





关键问题

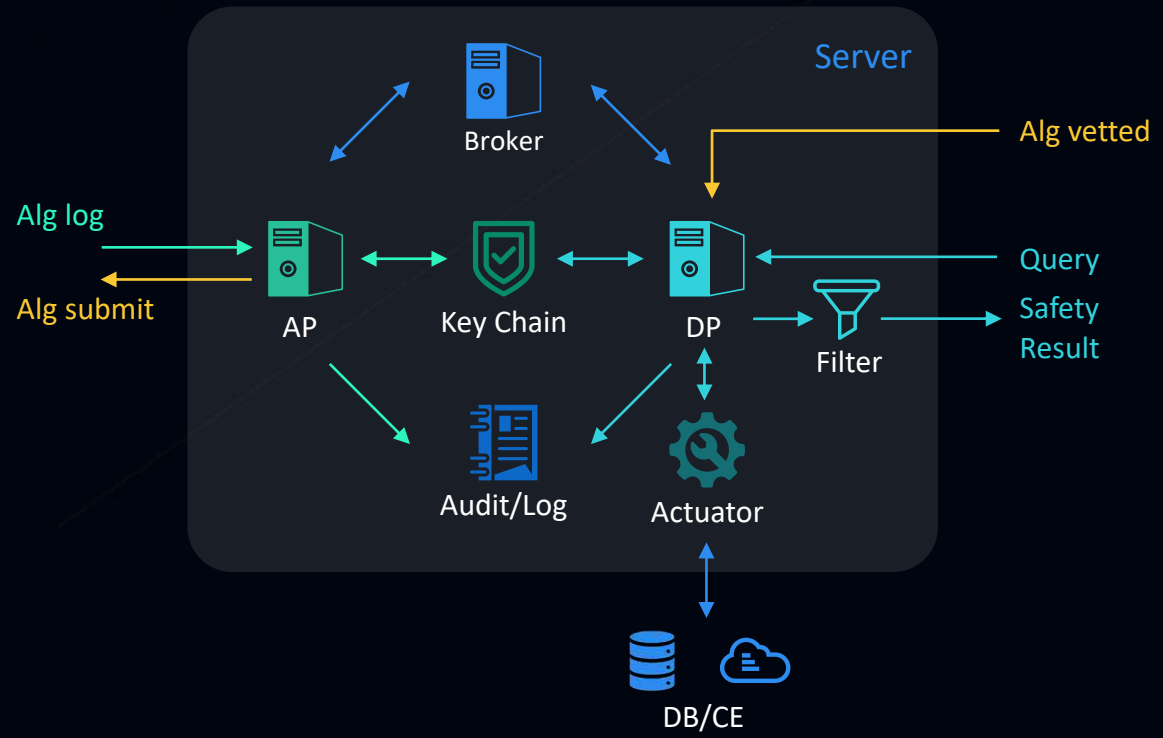
需求方如何
保护商业机密

需求方供应方如何
防止非授权数据泄露
满足法律法规要求

成本、效率以及
准确计量的问题

TalkingData的一些研究

OPAL



- 不经意传输
- 多方安全计算
- 联邦学习

解题思路

交付

由于需求方不能泄露信息，因此需要供应方交付全量数据。
为了防止非授权信息的访问，需要加密全量数据。

使用

需求方把选中的数据交给供应方解密，但为了防止供应方回溯主体标识，需要再进行一次解密，但是需要互不影响。

核心原理

同态加密的定义

业界

同态加密是一种加密形式，它允许人们对密文进行特定形式的代数运算得到仍然是加密的结果，将其解密所得到的结果与对明文进行同样的运算结果一样。

TD增强
同态性

如果一个加密算法，对它输出的密文先做计算再解密得到结果，与先解密再进行同样的计算得到的结果相同，则称该加密算法具有同态性。

TD的同态加密

假设有两个加密算法 E_1、E_2 以及对应的解密算法 D_1、D_2

使之满足对于经过 E_1和 E_2 顺序加密得到的密文 em=$E_2(E_1(m,k_1),k_2)$

如果能使 m=$D_1(D_2(em,k_2),k_1) = D_2(D_1(em,k_1),k_2)$ 成立

我们称这两个加密算法具有同态性

也就是说这两个算法的解密顺序可以交换

同态加密过程——交换律

数据供应商

企业客户

数据供应商

企业客户

数据形式

(D, S1*K1)

(S1*K1*K2)

(S1*K1*K2/K1) =(S1*K2)

(S1*K2/K2) =(S1)

方案说明

- 数据供应商对数据集中的标签信息统一使用K1加密
- 预置数据集R

  数据集 R
  (主体标识D, 标签信息 S1*K1)

- 客户从数据集中选取目标客群，并加密对应的标签数据，加密后返回数据供应商
- 客户返回标签数据
  (S1*K1*K2)

- 数据供应商解除之前自身加密，解密后返回客户
- 数据供应商返回数据
  (S1*K2)

- **客户对数据解密，得到最终标签数据结果**
- 数据标签结果S1

方案优势

供应商

- 无需泄露非查询主体信息

企业

- 无需输出任何主体信息

供应商

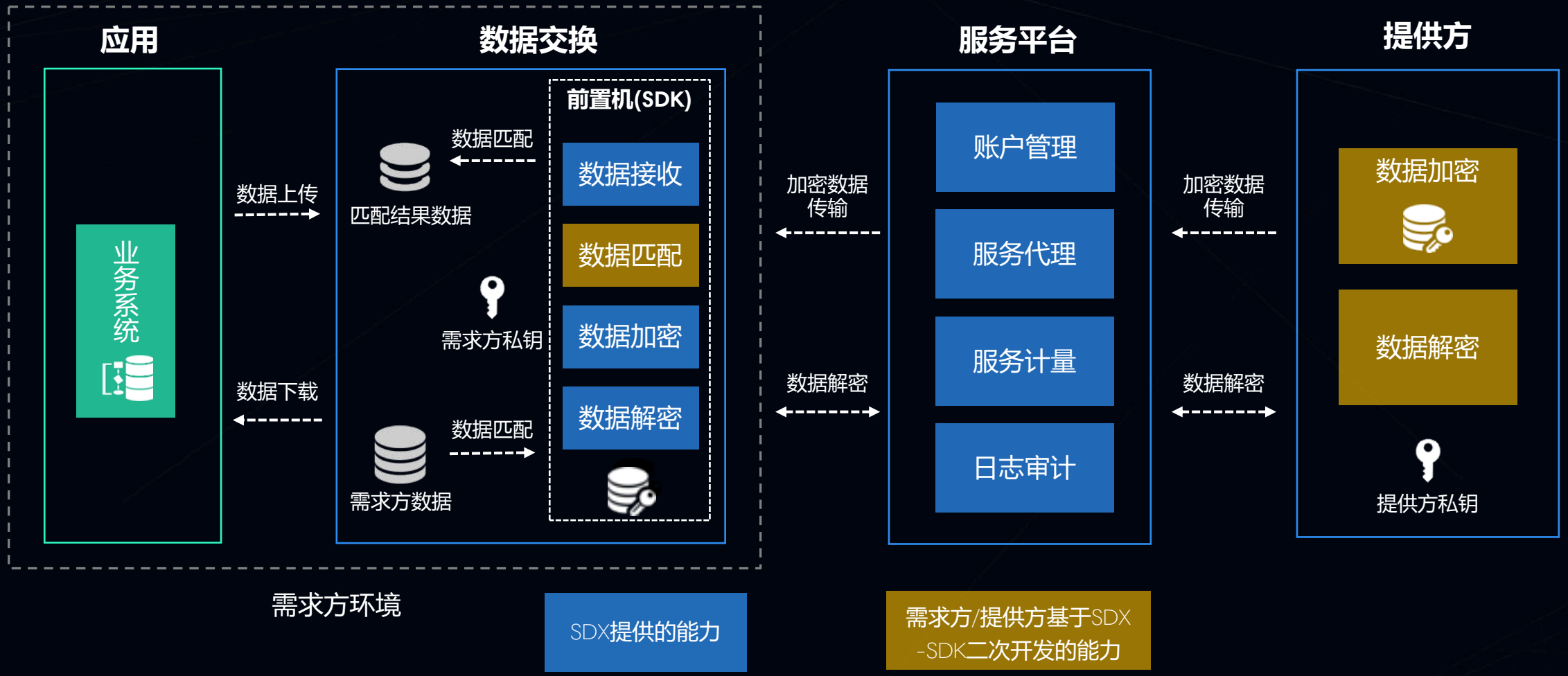
- 可统计实际查得数据标签数量

企业

- 安全查得标签信息

具体实现

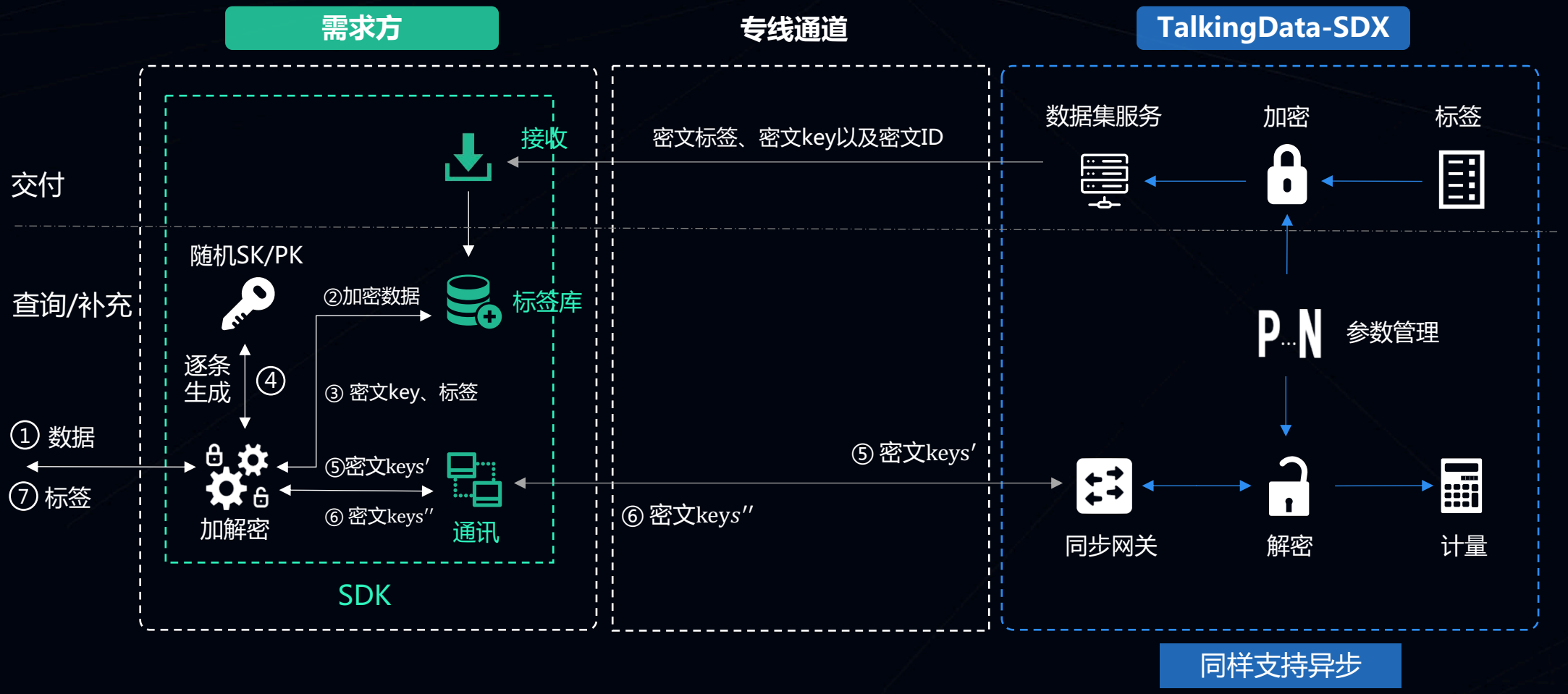
应用方案——SDX


应用
数据交换
服务平台
提供方
业务系统
数据上传
数据下载
前置机(SDK)
数据匹配
匹配结果数据
数据接收
数据匹配
数据加密
数据解密
需求方私钥
数据匹配
需求方数据
加密数据传输
数据解密
账户管理
服务代理
服务计量
日志审计
加密数据传输
数据解密
数据加密
数据解密
提供方私钥
需求方环境
SDX提供的能力
需求方/提供方基于SDX-SDK二次开发的能力

实现细节及创新

- **对称与非对称加密组合使用**
  - 利用对称加密以及随机密钥（Key）加密数据；
  - 对Key使用支持同态特性的非对称加密，用于传输（协商密钥），提高安全性，也保证计算的稳定性；
  - 双方无需交换公钥；
- **概率加密**
  - 使用随机密钥或随机数进行逐条加密，一次一密，提高安全性；
- **分组加密**
  - 对交付的数据集合理分区，每个分区使用不同的公私钥对，提高交付数据集的安全性；
- **数据集压缩**

关键组件及流程


需求方
专线通道
TalkingData-SDX
交付
查询/补充
接收
密文标签、密文key以及密文ID
数据集服务
加密
标签
随机SK/PK
②加密数据
标签库
逐条
生成
④
③ 密文key、标签
P…N
参数管理
① 数据
⑦ 标签
⑤密文keys′
⑥ 密文keys″
⑤ 密文keys′
⑥ 密文keys″
加解密
通讯
同步网关
解密
计量
SDK
同样支持异步

数据格式

Id	a	b	atts
0	T30rXSKbe9UzutHM2B7TgcoltQ	C4W1YiX6JvomII9E7jCy5Uz3Sw	VqD1zMqtXQLP6HaAoUC1+g
1	DO4CG3+1MhChE+v3L1Q7ar37kA	R06R/pRUP0Z02J0sv+V/3BdJ7A	eRDlR6smBfbih4Ip9dupsw
2	Ia0JfCSWqVy1+l1P9sMrRBZI9w	MFngwzTvyDlIaoEiVHRUNagt7g	qKtarU0HgnkrB8w9NJZ9cw
3	X7CRcJUGT672IyoN1RYcOQ8VVw	QoPCazhywtopu6q944Dbgr6y7g	JZi3qGpKIBrtBA/kEDSmXg
4	FyqSTiGIPNFFK2I7r3l5pKODgQ	VP3lgBfGDTApmC99yxaWPblmXA	1OW/738XtqQxir4C98F0kxKlxZka/dR10RyVrjLTdPiD/avIjZ0xLtLMOnNHTVDf
5	QzDBDQ3xIl0+6ijJ4eRVN9jr5A	YelSkBz3AecxDa6gUZNRZhpyBg	rNelfn3rPI8CQiijqE8dug
6	YvZt5L6sF+JeG/Jm69KKIkmR3g	DQkb3xSCKkV6KmpnvKpcW6/Zxg	QM0zw8ZbrStR6MEqZ9kmQvlMcv9oHK+cEJZDBMhbaPg
7	MfjC0tBL8n4UGAm/2SAAmMNg7g	Bkb++7rI6dNKck69RqSfR5jzhA	2puJokg084eGvdPpDu7OpQ
8	RHExf04iSCbVuYtaLxdAMBueQw	T5ayvIrDIY9WAdHYfHiK9CujiQ	/6XrTYGbyVwSj1Hyd9Uw3w
9	bIdIB4RmTuPjimF+xKZ1FdBCfw	W99y89/BeOvcFz01h+Bz+iMhSA	tlLqFFHDSU0bvg2CSPmrWNceICoQA+NFzyV6e6V+UW4

	用户	数据源
A	WrMJkO9PDKvGOG2a29pPEjOWDg	Ia0JfCSWqVy1+l1P9sMrRBZI9w
B	COT6HGawUAz/b/kgnVd6yWmkeg	MFngwzTvyDlIaoEiVHRUNagt7g

数据源的响应： G3eCdFPdWi5XHXq1WiVxdDK3Og

用户计算后得到密钥28015353104927604560862917525128270011最终解密得到属性“att5,att6”

特性

安全性极高

- 采用同态及概率加密算法，且通用算法
- 不泄露需求方用户ID
- 平台无法截获任何有用的数据
- 传输的数据都是加密的

按实际使用计量

- 在不知道请求方内容的情况下可对数据进行统计。

数据交换效率高

- 解密服务调用过程只传输加密后的密钥，大大减少网络传输数据量
- 单条数据处理时间固定
- 不产生其他额外的数据计算

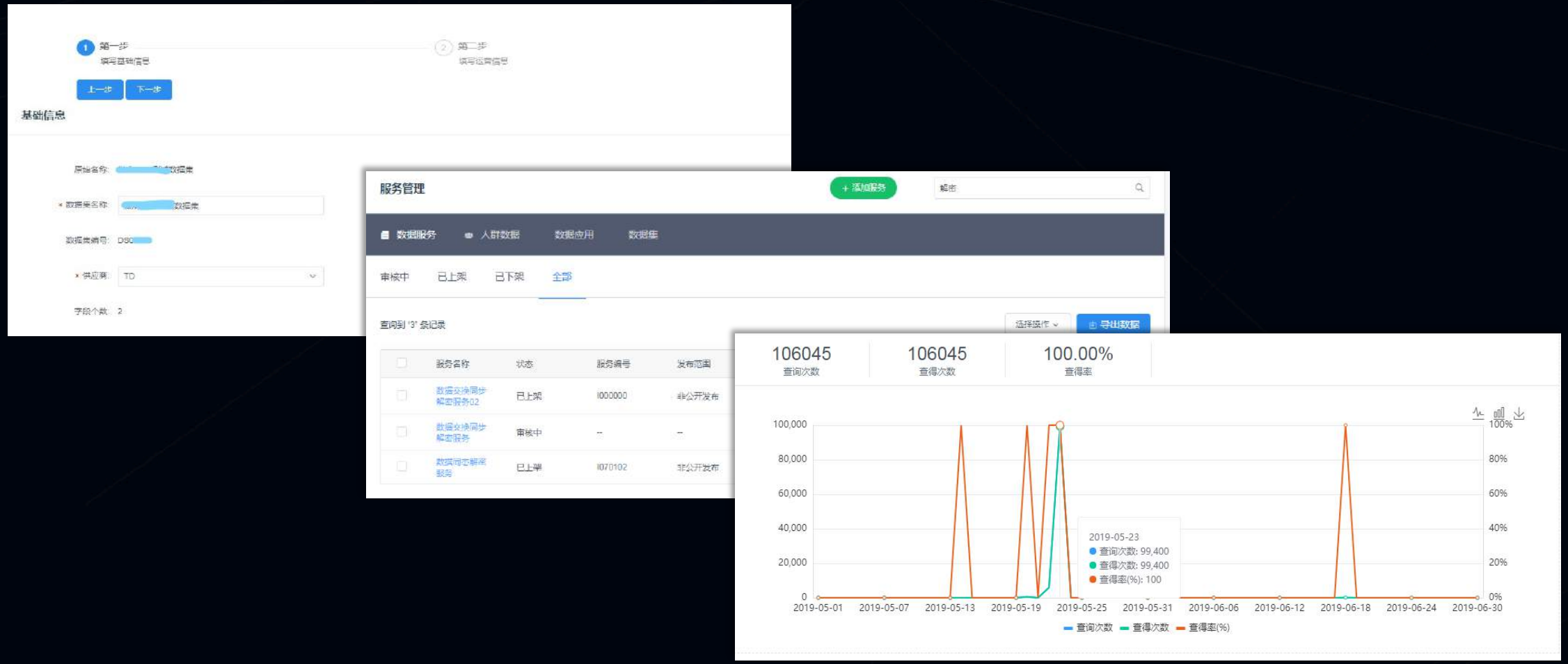
快速开发

- 提供SDK实现数据接收、匹配、解密等功能，可快速对接现有系统或进行定制开发

适用场景

查询或补充外部数据——单方使用

应用展示



TalkingData

应用展示

```
Last login: Wed Nov 20 10:17:55 on ttys005
java -cp /Users/victor/code/tdProjects/demoexcute/target/demo-excute-1.0-SNAPSHOT.jar com.talkingdata.exchange.demo.Upload ~/Desktop/sample-500.csv
win98:~ victor$ java -cp /Users/victor/code/tdProjects/demoexcute/target/demo-excute-1.0-SNAPSHOT.jar com.talkingdata.exchange.demo.Upload ~/Desktop/sample-500.csv
alluxio path:/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8
datasetId:16d1ff2e71f64bab94f7337af3a531bf
jobId:20efee415d3d42eeacb15448b40318ea
groupId:3a95fd4d9a404ceaa59e4fab020226f4
salt:Uegf7e8g7dcfahij8KowAA==
{"jobName":"test-data","jobstatus":10,"datasetId":"16d1ff2e71f64bab94f7337af3a531bf","startTime":1574219040,"updateTime":1574219040,"inputSettings":[{"inputId":"5fb39e0c54f444da861b97e6cfb655e2",
"path":"/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8","recordNumber":0,"storeId":"default-alluxio","taskId":"20efee415d3d42eeacb15448b40318ea"}],"outputSettings":[{"path":"/encrypt-result-2
0efee415d3d42eeacb15448b40318ea.zip","recordNumber":0,"outputId":"780c977632884e8684c230022ce2aa8c","runtimeParam":"{\"token\": \"10a56c36-f88a-43e2-bf6f-42704111c0cd\"}","storeId":"store-sdmk","
taskId":"20efee415d3d42eeacb15448b40318ea"}],"taskId":"20efee415d3d42eeacb15448b40318ea"}
{"jobName":"test-data","jobstatus":10,"datasetId":"16d1ff2e71f64bab94f7337af3a531bf","startTime":1574219040,"updateTime":1574219040,"inputSettings":[{"inputId":"5fb39e0c54f444da861b97e6cfb655e2",
"path":"/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8","recordNumber":0,"storeId":"default-alluxio","taskId":"20efee415d3d42eeacb15448b40318ea"}],"outputSettings":[{"path":"/encrypt-result-2
0efee415d3d42eeacb15448b40318ea.zip","recordNumber":0,"outputId":"780c977632884e8684c230022ce2aa8c","runtimeParam":"{\"token\": \"10a56c36-f88a-43e2-bf6f-42704111c0cd\"}","storeId":"store-sdmk","
taskId":"20efee415d3d42eeacb15448b40318ea"}],"taskId":"20efee415d3d42eeacb15448b40318ea"}
{"jobName":"test-data","jobstatus":10,"datasetId":"16d1ff2e71f64bab94f7337af3a531bf","startTime":1574219040,"updateTime":1574219040,"inputSettings":[{"inputId":"5fb39e0c54f444da861b97e6cfb655e2",
"path":"/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8","recordNumber":0,"storeId":"default-alluxio","taskId":"20efee415d3d42eeacb15448b40318ea"}],"outputSettings":[{"path":"/encrypt-result-2
0efee415d3d42eeacb15448b40318ea.zip","recordNumber":0,"outputId":"780c977632884e8684c230022ce2aa8c","runtimeParam":"{\"token\": \"10a56c36-f88a-43e2-bf6f-42704111c0cd\"}","storeId":"store-sdmk","
taskId":"20efee415d3d42eeacb15448b40318ea"}],"taskId":"20efee415d3d42eeacb15448b40318ea"}
{"jobName":"test-data","jobstatus":10,"datasetId":"16d1ff2e71f64bab94f7337af3a531bf","startTime":1574219040,"updateTime":1574219040,"inputSettings":[{"inputId":"5fb39e0c54f444da861b97e6cfb655e2",
"path":"/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8","recordNumber":0,"storeId":"default-alluxio","taskId":"20efee415d3d42eeacb15448b40318ea"}],"outputSettings":[{"path":"/encrypt-result-2
0efee415d3d42eeacb15448b40318ea.zip","recordNumber":0,"outputId":"780c977632884e8684c230022ce2aa8c","runtimeParam":"{\"token\": \"10a56c36-f88a-43e2-bf6f-42704111c0cd\"}","storeId":"store-sdmk","
taskId":"20efee415d3d42eeacb15448b40318ea"}],"taskId":"20efee415d3d42eeacb15448b40318ea"}
{"jobName":"test-data","jobstatus":30,"datasetId":"16d1ff2e71f64bab94f7337af3a531bf","startTime":1574219040,"updateTime":1574219064,"inputSettings":[{"inputId":"5fb39e0c54f444da861b97e6cfb655e2",
"path":"/sdx-upload-input-58ec6fc0c9a541e597b3e53be4f73ae8","recordNumber":500,"storeId":"default-alluxio","taskId":"20efee415d3d42eeacb15448b40318ea"}],"outputSettings":[{"path":"2_01_2019112011
0423_edb255","recordNumber":0,"outputId":"780c977632884e8684c230022ce2aa8c","runtimeParam":"{\"token\": \"10a56c36-f88a-43e2-bf6f-42704111c0cd\"}","storeId":"store-sdmk","taskId":"20efee415d3d42e
eacb15448b40318ea"}],"taskId":"20efee415d3d42eeacb15448b40318ea"}
sdmk fileId:2_01_20191120110423_edb255
```

性能指标

QPS

5000/节点

20C/40T

处理时间

毫秒级

数据膨胀

- 毫秒级1024位密钥：344B/条
- 2048位密钥：688B/条
- HASH校验（可选）：64B/条

回顾与展望

回 顾

- 面临的问题
- 对同态加密的创新
- 实现中的增强
- 产品的特点、适用的场景

展 望

- 国密支持
- 减少数据膨胀
- 压缩率


谢谢
2019
TalkingData