**“公司业务数据环境”过渡性调整方案**

考虑到“兰州项目实际实现”和之前讨论确定的“总体技术方案1.0”的差异依然较大，而且估计至少半年内还会沿用当前计算和同步方式，所以为了解决数据删除的同步问题和其他潜在问题，需要重新设计合适的方案。

一、短期调整

总体思路是App\_Data库不变，新增Mongo库（假定叫M2）供应用环境（和兰州项目相关），同时数据平台组提供App\_Data到M2的单向数据同步程序。具体包括：

|  |  |
| --- | --- |
| **事项** | **描述** |
| App\_Data的不变 | 主体结构不变不影响原有大部分程序的访问。 |
| App\_Data的变化 | 1、App\_Data内添加“记录变动表”，主要字段是主键、表名、object\_id、site\_record\_id、删除时间。  2、所有数据新增要修改\_in\_time和\_utime字段；所有修改要修改\_utime字段；所有删除要记录到“记录变动表”中。短期内应该不会新增、删除表。  3、site\_record\_id的原有生成逻辑有严重缺陷，会作统一调整，但会考虑实现值的延续性（减少或消除对下游的影响）。 |
| M2相对App\_Data的不同 | M2是采用逻辑删除的思路。所以没有“记录变动表”。 |
| App\_Data->M2的同步程序 | 将App\_Data的增删改行为映射到M2的新增、逻辑删除、修改行为上。 |

思路本质：

数据分层、为中期调整做准备。

不解决的问题：

App\_Data很多同事或程序有访问权限，有隐患。

不解决记录误删除的问题。

不解决新增、删除表的问题。

其他注意事项：

数据同步是周期性行为，时间、频度需要商讨。

M2会占用许多存储空间。

二、中期调整

在上述数据分层基础上，会进行如下调整（但不影响M2的使用）。

1、App\_Data会逐步弱化和废弃，换用HBase做数据仓库。存储内容为全量时间维度上的数据，M2应该只需最新时间点的数据。

2、数据仓库会进行各层级的标准化工作，并完成数据的各种清洗和融合操作。

3、原“App\_Data->M2的同步程序”会做调整，但是会保持M2写入数据的兼容性。