**1、一个项目有多少张事实表和维度表？做过的宽表有多宽？**

答：大概3、4张事实表，维度表就比较多了，宽表有多宽的话，看具体情况，比较款的时候大概关联十几张表吧。

这个问题不要直接说死了 ,不然会很奇怪。

你要先说一下 你讲的是哪个项目 ，比如最近的项目 做了多久 ，然后 做了多少个版本 ，然后说最近一个版本 你负责的是哪一块。比如数仓，你负责了3个需求 ，创建了三个实事表和2个维度表 。

宽表一般是实事和维度表关联以后的大表（实事表和维度表横向拼接的），这个是集市层的表。

如果你又说实事和维度表 ，又说宽表就比较奇怪。

**2、你们数据库都是用同一个账号吗？**

答：不是的，obs、dw、dm、 和测试人员各用一个账户。

开发环境是同一个开发账号。除非你们的数仓不是同一个库。

测试环境是找测试同事申请的 查询权限的账号。

**3、上线跑了一下数据，应该有220条数据，结果只有200条，按照你的经验，你会怎么做？**

答：不会

|  |
| --- |
| 数据量变少了 分两种情况，  第一种，多次提交导致数据跑完以后 可能出现 部分成功 部分失败 。  比如前面的数据跑成功了 ，但是后面的数据跑失败了 ，这时候只是把前面的数据提交了，（这种问题的原因就是每次做 INSERT 或者 UPDATE 的时候 都做一次 COMMIT ），导致正常的数据可以插入，一旦数据同步失败后面的就不插入了，这时候要去看日志。  另外一种情况，最后进行 COMMIT  这种做法的好处就是 要么都成功 要么都失败，就不会出现部分成功 部分失败的情况，  在这种情况下 如果数据量少了 ，大多数是取数逻辑问题。 |

**4、你用的ETL工具是怎么调度的**

答：我们用的是工具是kettle，转换里面新建 输入 和输出 ，shift键连接起来，然后在输入和输出里面分别编辑步骤，分别连接各自的数据库，连接好了测试一下，是否成功。完事之后可以预跑一下，没问题的话，就可以启动了。后面可以创建作业，就是将之前的转换以定时执行的方式运行。

|  |
| --- |
| 你可以说是通过 CTM 调度的 ，不要说用 KETTLE 调度，我们一般很少用 KETTLE 做调度工具，但是可以用 KETTLE 执行存储过程。 |

**5、你觉得DW层的作用是什么？**

答：不会

|  |
| --- |
| 数据汇总（新增数据 和 历史数据）， 根据业务指标计算的需要进行建模 ，方便 DM层取数。 |

**6、上线后运维多久找我们一次？你的代码质量如何？**

答：

|  |
| --- |
| 直接说这个我没仔细统计过，我们后台的BUG很少主要是前端展示的问题，有时候我也会协助其他同事一起定位问题。  我们项目组的代码质量还不错，大多数问题都在测试阶段发现并结局了，上线以后的问题倒是不多 。  **总之 逻辑上没问题，如果非要有问题就是 性能问题 。** |