

### 文章

[Nicky Zhu](#) · 五月 7, 2022 阅读大约需 3 分钟

## 一个新库引发的血案

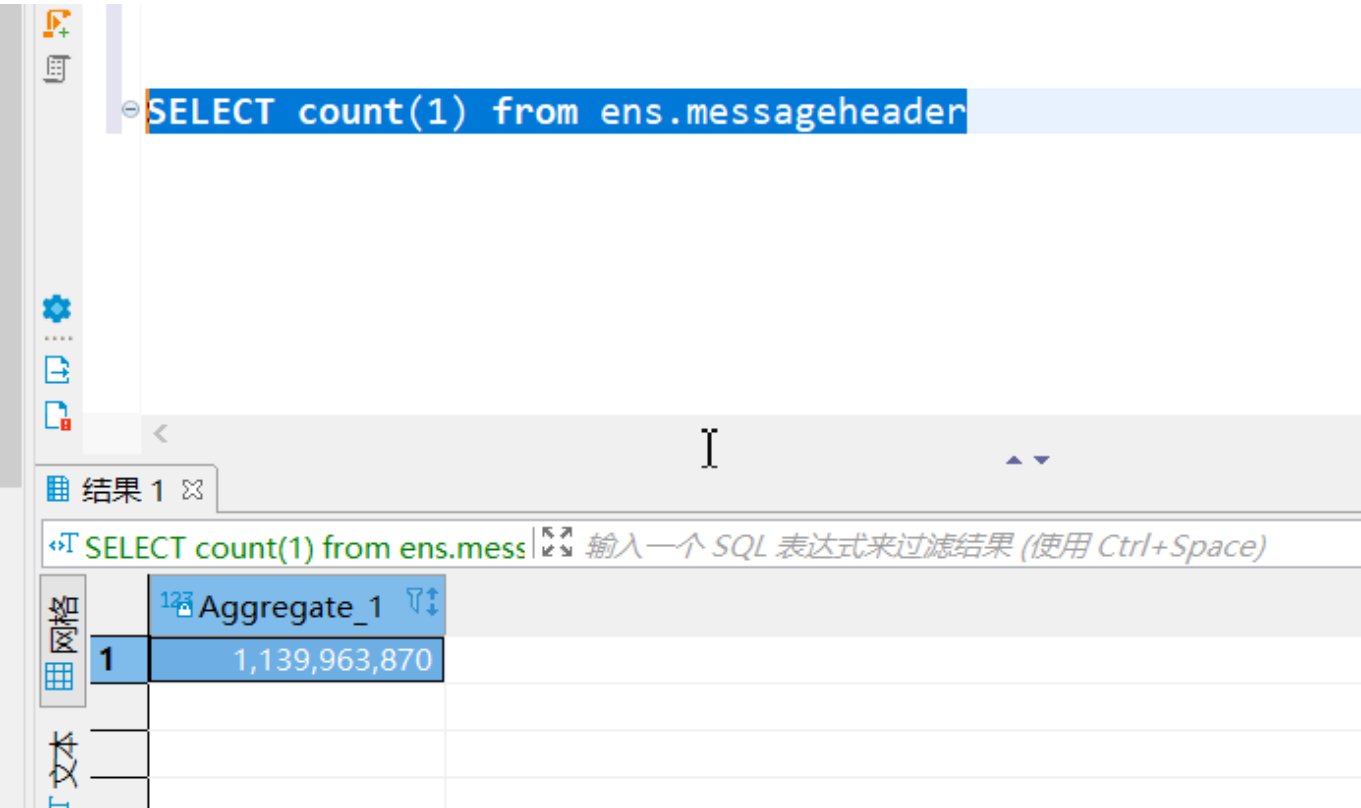
我们的一位客户五一期间向使用IRIS搭建的数据流推送一家三甲医院数年的历史数据，导致实施的同事们经历了一系列噩梦，包括但不限于：

1. 由于未通知实施团队有这样规模的数据推送，数据推送过程与全库备份任务重叠。尽管实例和数据流正常运行，但备份任务与数据流争抢IO，导致备份任务不能在预期时间内完成，实施童鞋五一加班处理问题。
2. 为了节省磁盘空间，服务器上部署了定期删除IRIS备份文件的任务，原本能够保持一周的全备+增量备份，但在本次数据暴增的情况下，新的备份尚未完成而旧的全备已被删除，导致问题发生时没有可用于恢复的备份。
3. 由于这次数据推送前未进行数据质量校验，推送的数据全部不合规，但已经历了较长的数据流进行处理全部入库；同时由于备份文件已被删除，无法通过恢复数据库的方法回滚，导致实施童鞋不得不逐条从生产环境三个库的数百张表中挑出问题数据逐一删除，从五一放假结束至今还未完成善后工作。大家可以设想一下，如果备份还在，那么恢复备份就可以了。

因此，我们希望再次提醒各位在前线奋斗的亲们：

1. 善待你的备份。尽管对于大型医院或医疗集团来说，两周的全备+增量备份策略下，备份文件会占据数个TB的存储空间。但在需要回滚时，这几个T的空间能救命。
2. 保持可用的测试环境。尤其是对于可能出现随机数据需求的客户，随机产生数据需求意味着随机出现测试需求。
3. 验证新数据的合规性，永远不要假设新数据一定合规。未经测试的新数据必然毫无悬念地导致新问题。
4. 对于任何批量数据处理任务，请务必提前规划，错开资源（CPU、内存、IO）的抢占，避免抢不到资源的任务饿饭。
5. 保持与最终客户的频繁沟通，所有对于生产环境进行的改动都应该经过项目组评估。虽然客户是上帝，但命运有时很顽皮，生产环境的安全保障也需要客户的合作。

最后，大家都知道InterSystems的IRIS在多数客户的场景下都不需要搭建负载均衡集群，这家客户也不例外，数据流中的数层结点上部署的都是单实例IRIS，通过Mirror实现高可用。在这次新数据的上传过程中，IRIS的数据流自然经历了突如其来的爆发式数据压力，以其中一个实例的消息量为例：



该用户在实例上保存30天的数据，可见在经历了五一的消息暴增之后，该客户的每日平均消息量已超过3300万条每天（实际上我们已经查到其中数天单日消息增量已超过5000万条），而该客户平时的消息量不过数十万条每天。

这次IRIS经洪峰而不倒固然可喜可贺，但相信在需要在客户面前经历各种千夫所指的PM、实施、开发与测试同事一定不希望经历这种惊喜。

Good luck.

[#InterSystems IRIS for Health](#)

源  
URL:  
<https://cn.community.intersystems.com/post/%E4%B8%80%E4%B8%AA%E6%96%B0%E5%BA%93%E5%BC%95%E5%8F%91%E7%9A%84%E8%A1%80%E6%A1%88>