## Elasticsearch用作日志分析中的数据存储以及搜索的考量

### 不适合的理由：

1. **mapping不可改，不能改index属性。**ES中以定义的mapping不能修改名字和属性，无法修改名字勉强还能接受，但无法修改属性。官方文档中介绍了几种修改mapping的方法。一个是新建一个字段，程序中所有地方修改名字，这对于复杂的项目容易出错，而且无法保留原来的数据；另一个是利用alias创建一个新的索引，但是所有数据需要重新导入，这需要很长时间，操作性不强。
2. **无法多对多。**ES中提供3种关联关系：Field collapsing（严格来说不是关联），Nested object，Parent-child。前两种都是直接将一个mapping声明在另一个mapping中，第三种关联是在创建子文档是指明他的父文档，但是一个子文档只能有一个父文档，因此也不能实现多对多的关联。其实如果理解了ES的目的是提升检索效率，就不难理解为什么没有多对多关联了，在关系数据库里这就是个效率瓶颈。
3. **没有用户验证和权限控制。**ES本身的访问权限可以通过nginx进行控制，但是同一个ES中不同索引间目前是没有权限控制的。
4. **从ES设计的初衷看，为了检索，为了统计。**这个从字段的store属性中可以看出来，查看ES手册（https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-store.html）可以发现，默认情况下字段的原始值是不会被保存的，这跟数据存储是南辕北辙了。
5. **项目开始时不好确定shards数量。**少了以后扩展不方便，多了一开始影响性能。这个可以通过将type命名为doctype-yyyymmdd来解决，每天都生成新的一个或多个shard，但是注意在搜索时需要在doctype-\*中搜索。
6. **ES非常适合特定的需求，但不适合用于数据存储。**ES索引速度快，扩展方便，性能优异，但在功能上不适合作为数据库使用。数据存储的目的是为了以后能方便的使用，不仅是针对当前的需求，也要为未来可能出现的需求做准备。由于ES有以上几点问题，无法适应需求变化。

### 适用的场景：

1. 检索。ES本身作为一个搜索引擎，用来处理检索的任务再合适不过。你可以在线上项目中直接将内容写入ES以提供检索服务，也可以把以往的数据导入ES以处理特定的需求。关于ES和Solr的比较以后有时间的话会写一篇
2. 统计。ES的统计也是基于检索功能的，聚合功能使得统计结果处理起来非常方便。如果你只需要统计而不用检索，可能有其他工具更适合你，比如Spark SQL。

（这是一方面，重点关注不适用的场景）

### 知乎上对hbase和elasticsearch的对比

1. hbase面向列非常好加字段的！  
   es适合搜索和分析小规模数据，速度快过hbase。  
   hbase稳定可靠，而且可以通过mr spark等大批量拉取数据。
2. 从基本功能来说这两个确实有相似性，但是根据业务需求不同，我觉得有几点可以考虑：  
   1. 查询复杂度：HBase支持简单的行或者range查询，比如给一个PK查该行的数据，或者给一个begin/end查这个范围的数据，如果想完成更复杂的功能就不太容易。而ES支持的查询比较丰富，或者说这些查询都带有一点复杂计算的味道了。比如你有个论坛，你想查帖子里面是否包含敏感词，如果采用HBase就比较麻烦，使用HBase你可以将帖子存进来、读出去，但是要查内容里面的东西，只能一点点过滤；而ES是可以比较方便的帮助你完成这个功能的；  
   2. 数据量：按道理说两者都是支持海量数据的，但是据我个人感觉，HBase可能更容易支持更多的数据，因为其一开始设计就是解决海量问题的；而ES是后来慢慢增强其存储扩展性的；那么也就是说，HBase上手起来扩展性不太会阻碍你使用；ES可能要多费点劲。当然，听说也有人写了ES基于Azure或者S3的存储插件，但是稳定性不知道如何；  
   3. 剩下的就是比较远的考虑，比如维护性，HBase基于Hadoop那一套，组件多，维护起来代价也不低，而ES自成体系，维护起来稍微好点；当然这个是相对的，绝对来说都不会容易。比如新功能开发，比如成本控制等等。。。
3. ES 本质上是search engine。如果你的需求是简单的字符匹配或者统计数据的话，ES更合适。而且UI 可视化 tool 也很成熟

Hbase 本质上还是数据库，适合和HADOOP stack 一起使用，支持更强大的批量计算，流水日志有点over kill。

1. hbase存多读少，不适合高并发查询，适合存数据；  
   es是全文检索，适合日志分析日志统计之类。
2. <https://www.zhihu.com/question/45510463>