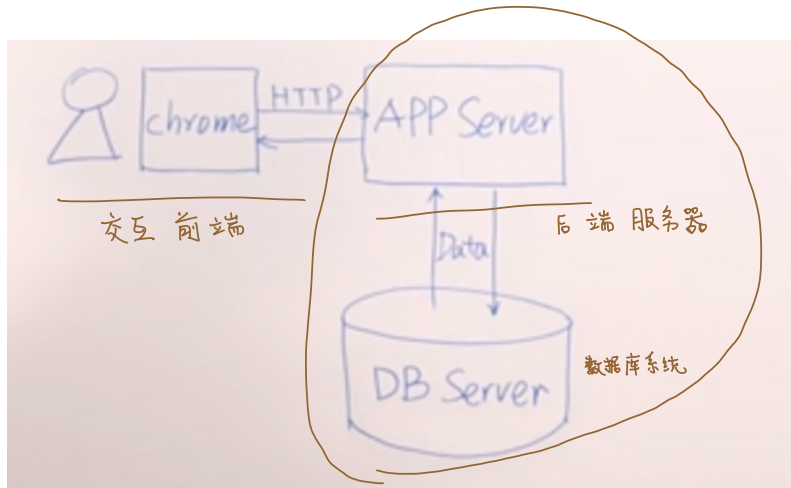
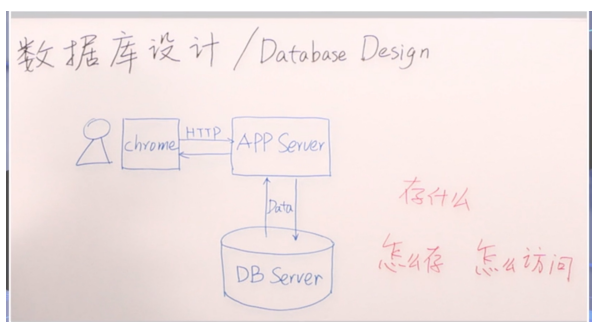


数据库设计的基本概念



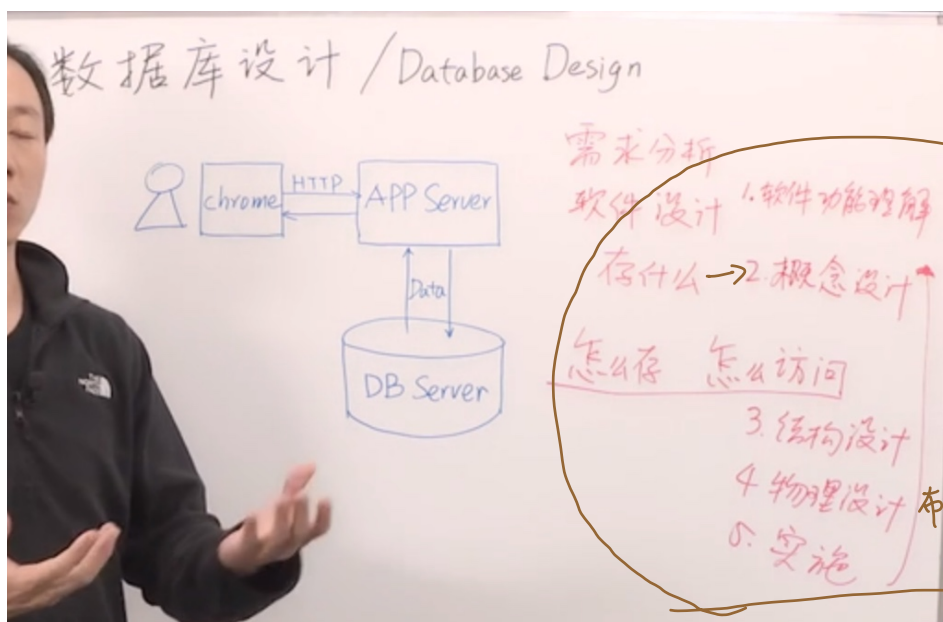
一个应用软件的三个部分。

存什么？怎么存？应用程序怎么访问数据？



需求分析

软件设计



需求分析

软件设计

存什么 → 2. 概念设计

怎么存 怎么访问

3. 结构设计

4. 物理设计

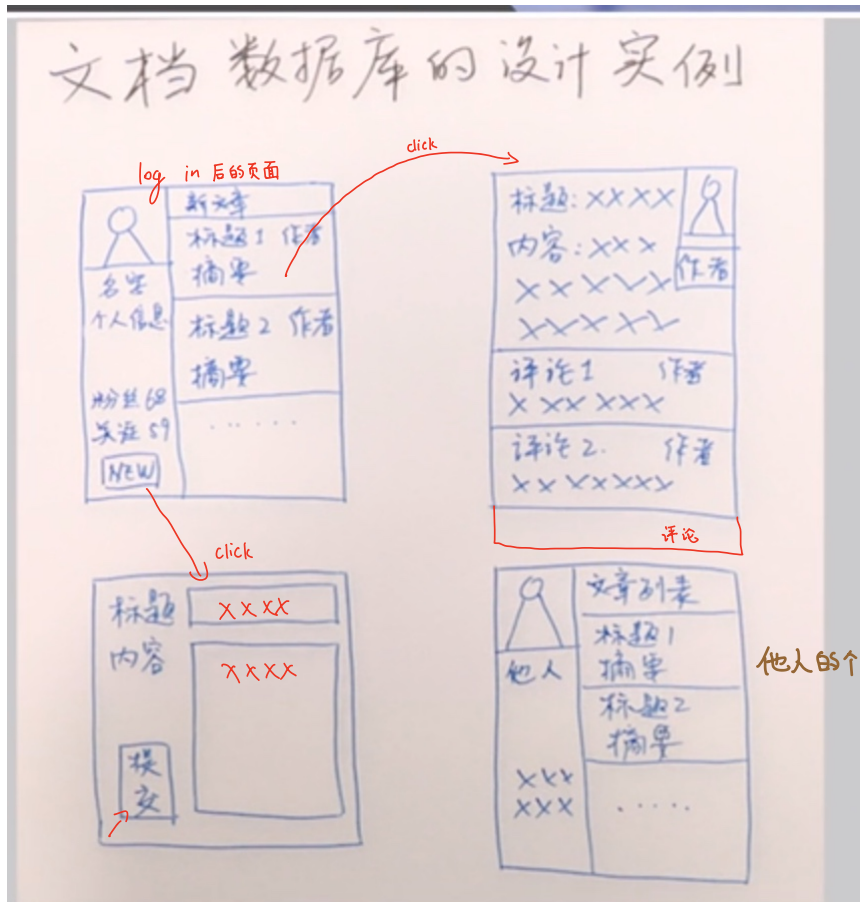
5. 实施

循环
迭代的过程

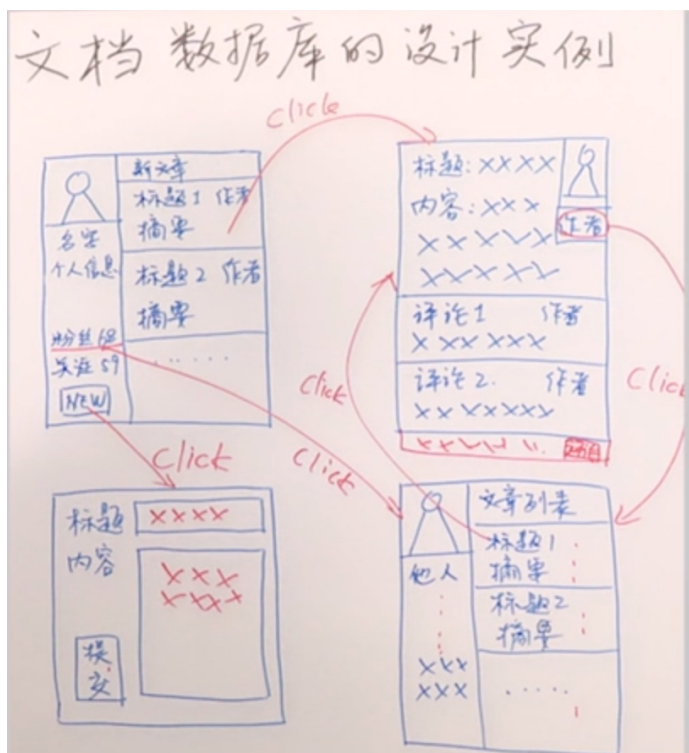
部署方式 / 索引

文档数据库设计实例：需求分析

blog

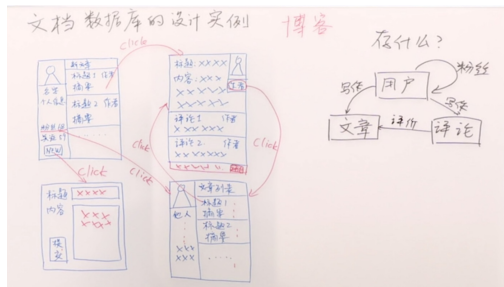


他人的个人主页



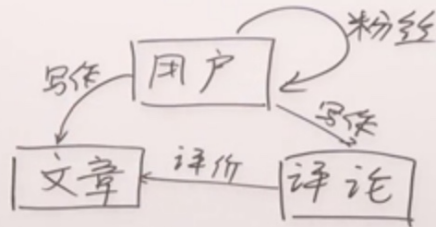
文档数据库设计实例：概念设计

存什么？



博客

存什么？



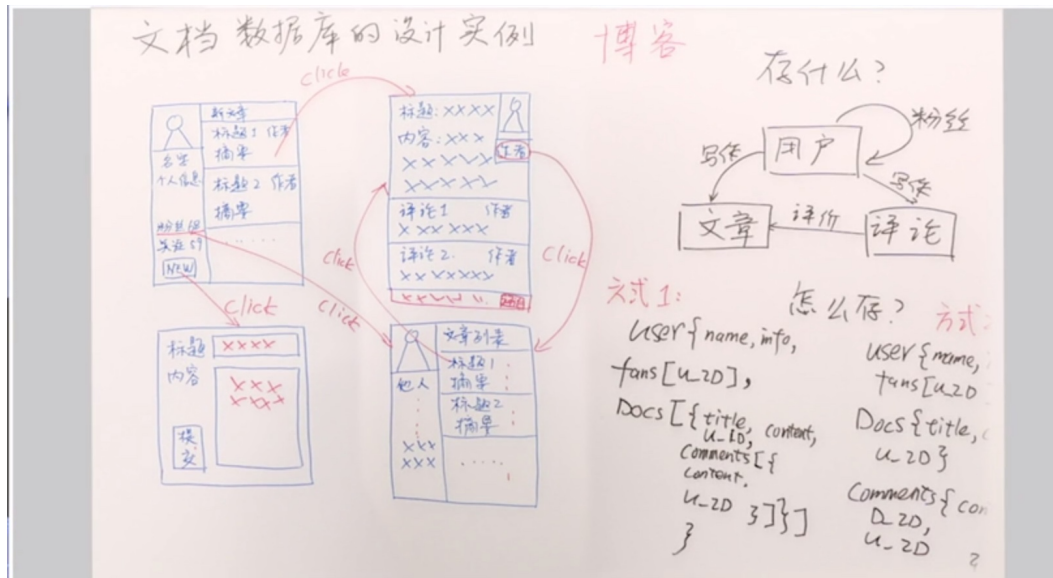
概念模型

怎么存？

```
user { name, info,
  fans [ u-2D ],
  docs [ { title, content,
    comments [ {
      content,
      u-2D
    } ]
  } ] }
```

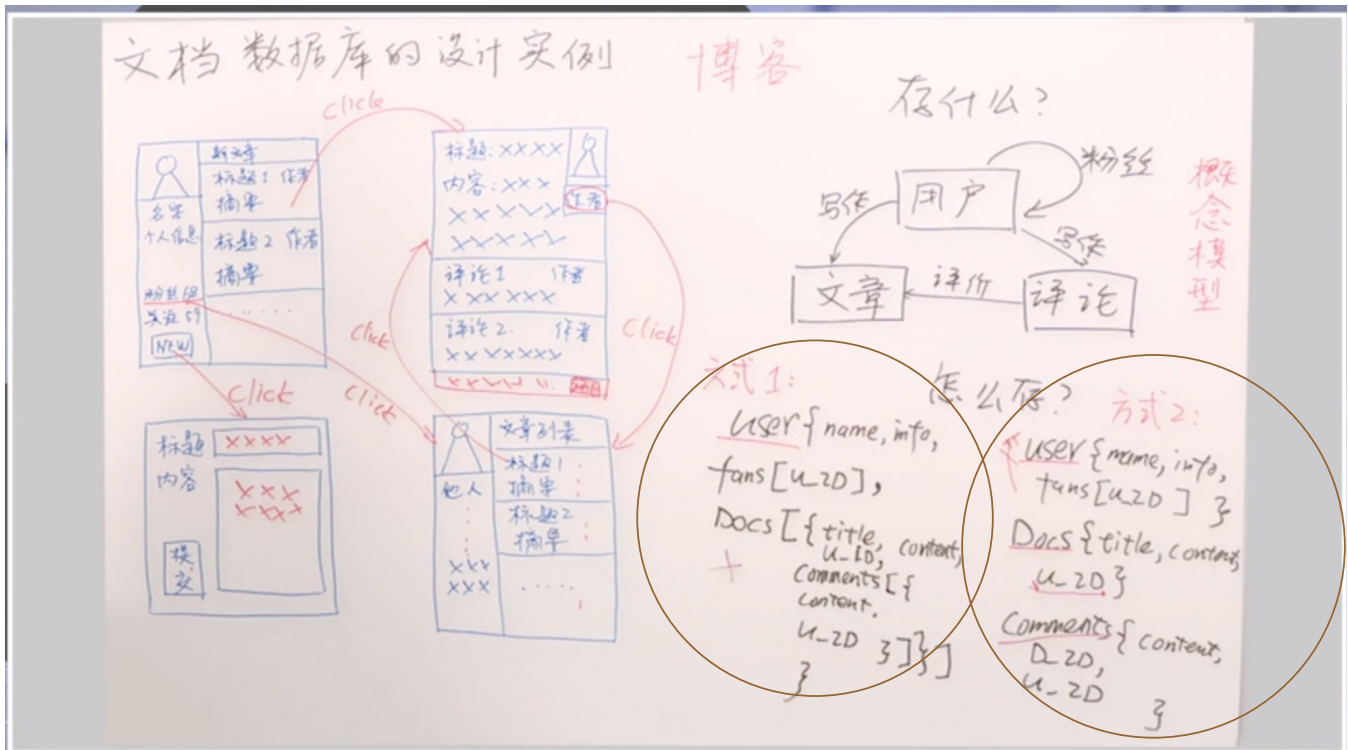
用户粉丝的文档 ID, 起指针作用

这个文档集合的每一个文档是关于一个用户的

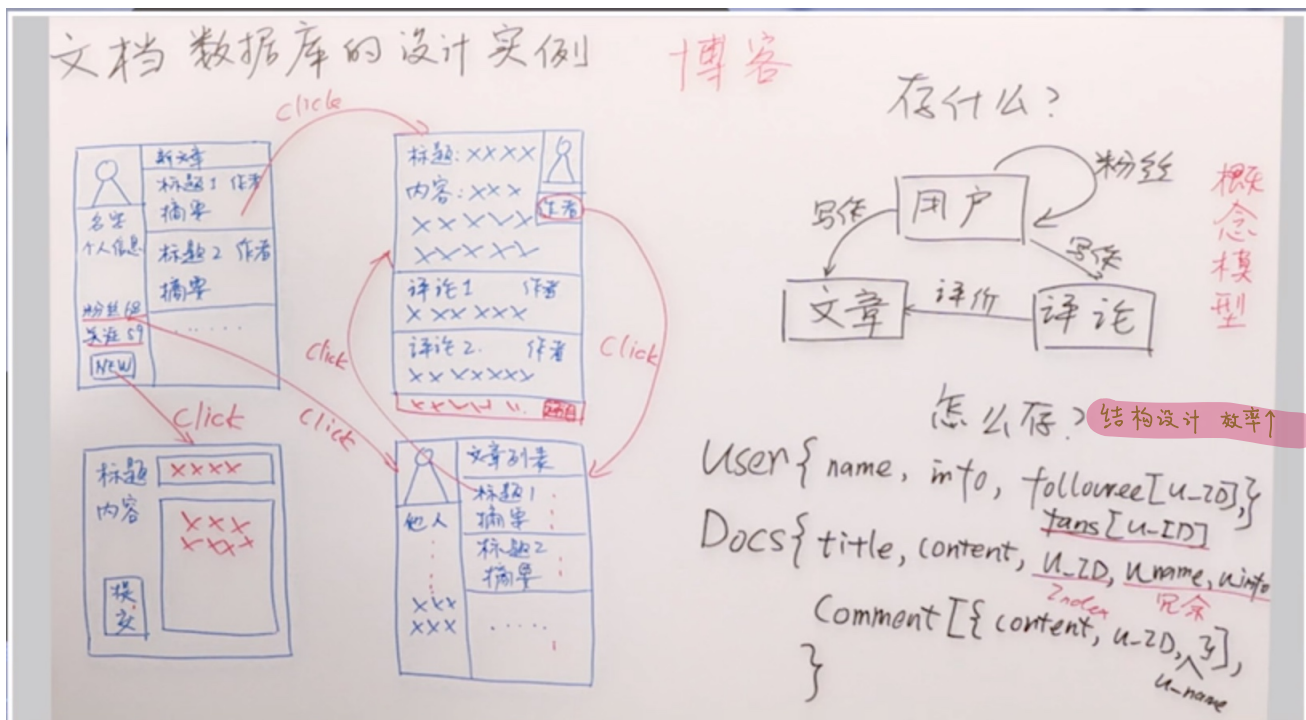


三个文档集合，每一个对应一种对象

不同文档结构设计的比较

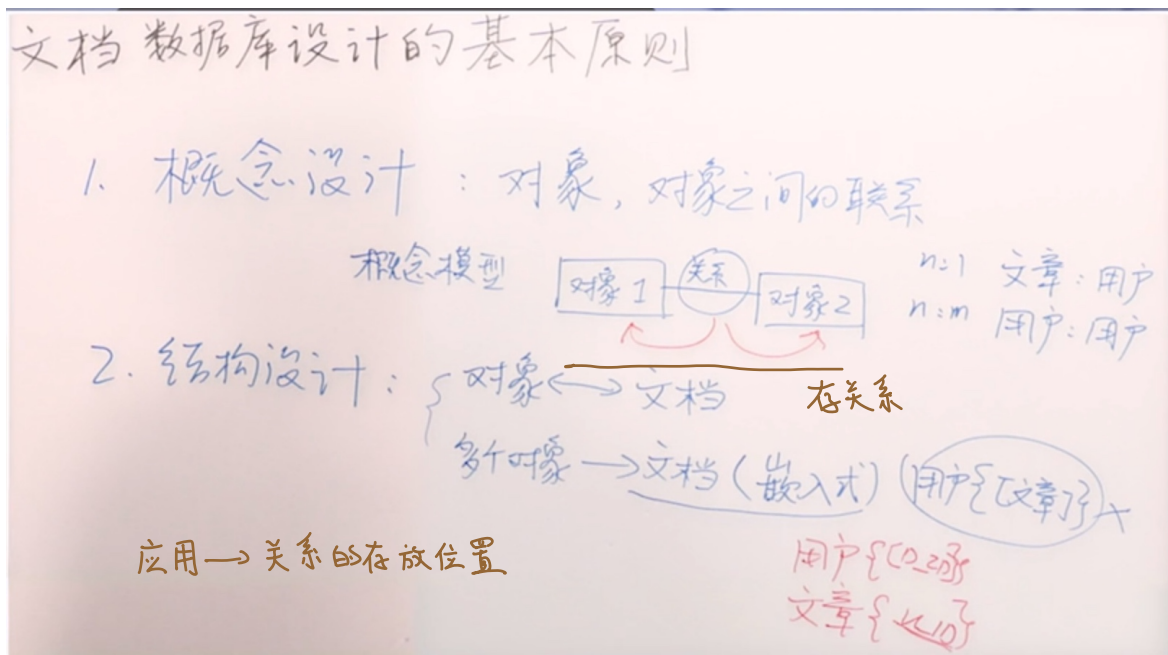


文档数据库设计实例：结构设计



文档数据库设计的方法 (1)

我一直会访问文章的同时访问他的评论，而不会脱离文章单独访问他的评论，所以评论适合嵌入在文章的文档里面。而用户和文章的关系，我会单独访问一个用户们也会单独访问一个文章，所以不适合嵌入。如果用嵌入式的话对象和对象的关系就一目了然了，比如说用户就是文章的作者。关系的存放位置根据应用的需求来定。

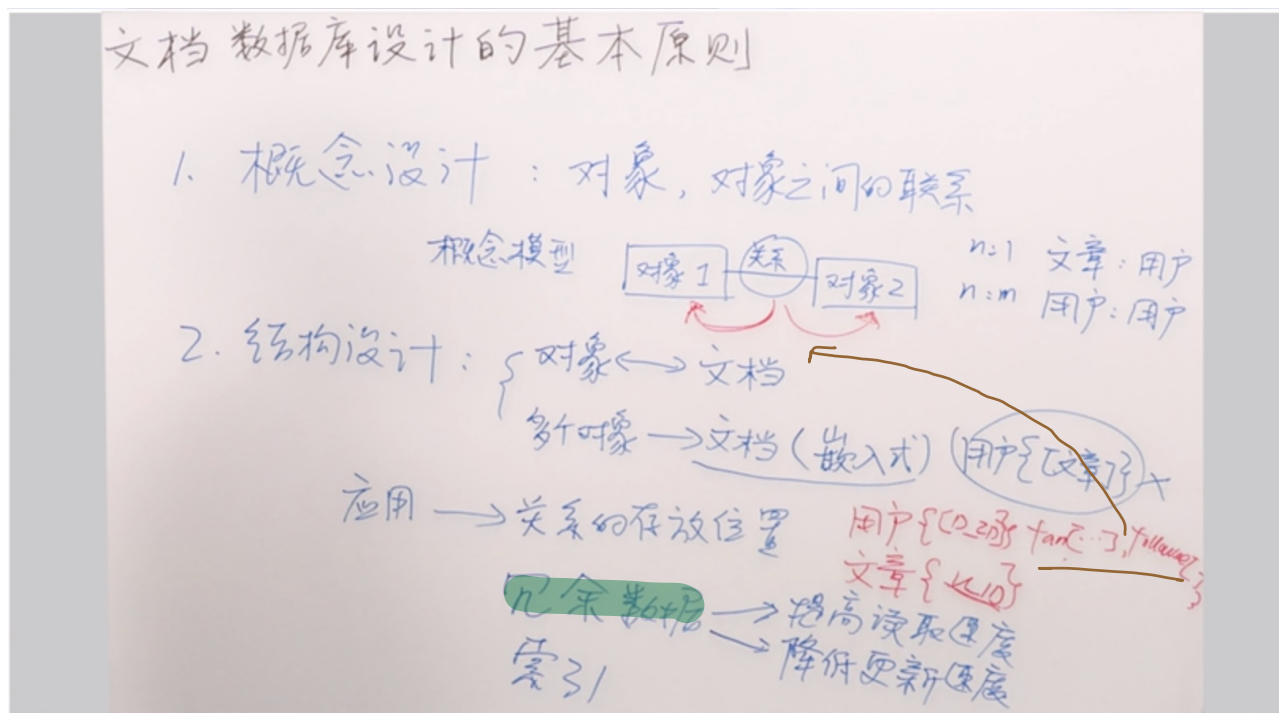


文档数据库的设计方法 (2)

冗余数据：用户的名字除了在用户的文档里还在文章的文档里，好处是增加读取数据的效率，但是有可能降低更新的效率（如果用户改名了）。

什么时候使用冗余，什么时候避免使用冗余呢？

看数据经常被读取还是经常被更新，经常被更新就不适合用冗余，基本不改动就可以用冗余，这样访问可以更直接。



不经常被改动但是经常使用的属性，比如名字，适合做冗余。如果一个属性经常被修改还做了冗余的话，更新代价会很大