# 短连接需求文档

## 需求背景

在一些业务中，例如业务推广的短信、活动短信、防止参数暴露等，需要将一些正常的较长的域名替换成一个短域名，也可以便于发布、传播、复制和存储。。

## 需求分析

需要生成短连接的服务，传入长连接，我们的服务给他返回一个合适的短连接，该链接的要求：

1. 短域名长度最大为 8 个字符
2. 短域名最好是字母和数字，可区分大小写
3. 短域名需要维持全局唯一

资料调研：

一种是基于MD5码的，一种是基于自增序列的。

1、基于MD5码 ： 这种算法计算的短网址长度一般是5位或者6位，计算过程中可能出现碰撞（概率很小），可表达的url数量为62的5次方或6次方。猜测google、微博用的是类似这种的算法，可能看起来比较美观。

2、基于自增序列 ： 这种算法实现比较简单，碰撞的可能性为0，可表达的URL可达无穷大，长度从1开始。

## 3、需求思考点

1、根据长域名生成短域名的算法

2、短域名需要全局唯一

3、长域名和短域名需要形成一对一映射

解决方案：

1、利用MD5算法，先将长域名进行前后随机加密，再进行MD5，生成一个32位的字符串A，然后再将A分成4段，每段8位B，再将B进行移位操作，生成一个长整型C，C在与0x0000003D进行“&”操作生成元数据的下标Index，使用该下标获取相应元素，循环组成相应长度的短连接。

2、使用Caffeine作为缓存，既可以解决内存溢出问题，还可以解决短域名全局唯一问题，同时形成一对一映射

3、此方法不借助外部缓存和组件可以实现高效的短域名服务

## 4、正常短域名一般在4-8个字符串

当短域名越短的情况下，碰撞就会越高，因此较短的短域名只适合小型的系统，也就是对短域名需求不是特别高，或者淘汰特别快的系统，因此中大型系统建议短域名在6到8位。