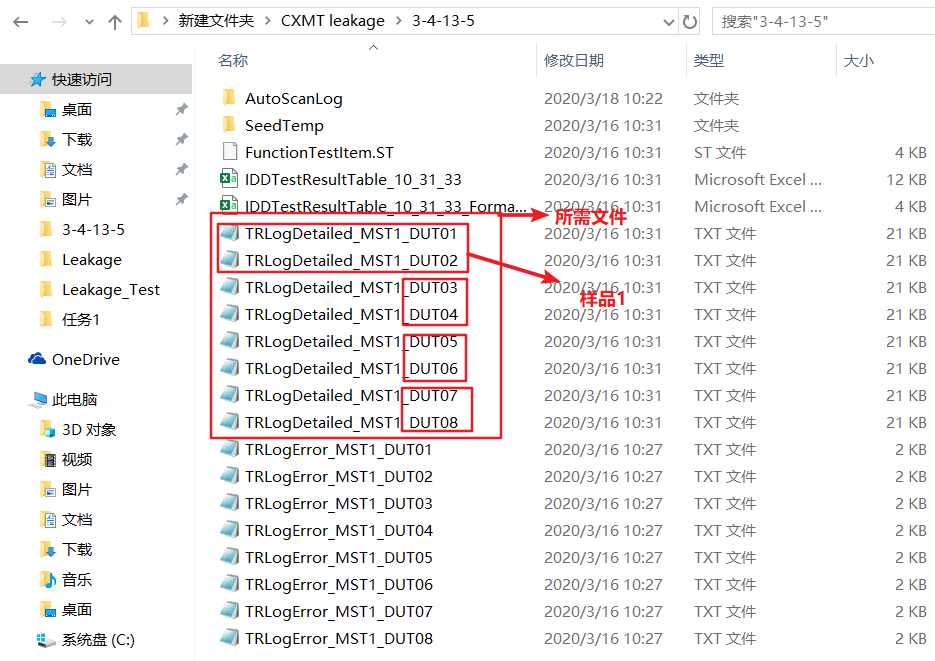
# Leakage数据提取与处理设计方案

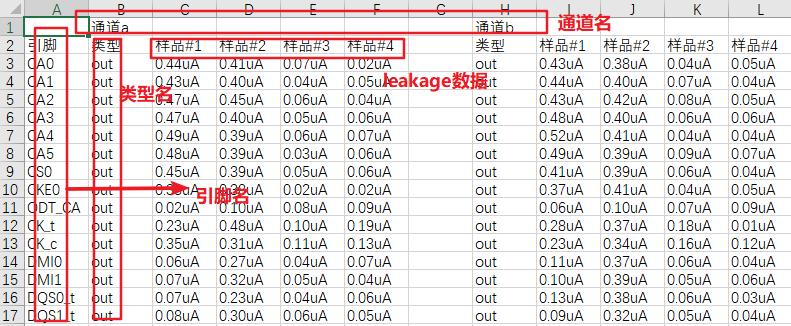
1. **题目概述**

现给定一个CXMT芯片的测试文件夹，需要从文档中多个测试log文本内提取芯片各引脚的Leakage数据。按照以下格式：芯片引脚名，通道名，输入输出类型，样品名及Leakage数值将其排列好并输出到Excel表格1中，再将所提取的数据按Leakage数值从高到底顺序进行排列，再次输出到Excel表格2中。

1. **题目分析**
2. **文本名称识别**

如下图所示，文件夹中有很多测试文本，而其中我们只需要前缀为TRLogDetailed\_MST1xx的测试文本。这8个测试文本的相邻的单双数DUT别名对应芯片b和a通道。需要判断文本名称，将文本内容中数据归类到对应的通道表格及样品表格。

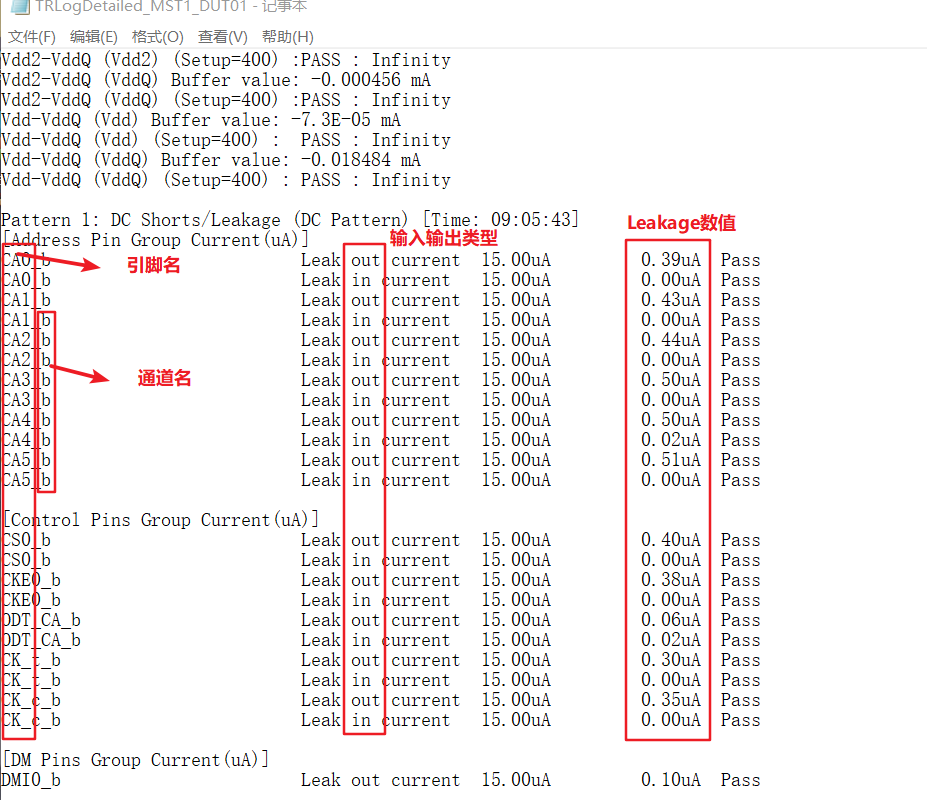




1. **字符串的提取**

在每一个目标测试文本中，我们只需要提取跟Leakage有关的有效信息,如图所示：

需要对文本内容进行扫描筛选出目标字符串，并将所搜索提取的信息进行分类，按照引脚名、通道名、输入输出类型、电流数值进行分类。将同一个引脚的所有详细信息打包成结构体，供后续数据处理。



1. **输出Excel表格**

对提取后的数据进行输出生成excel表格，可以生成.csv文件。csv文件称为逗号分隔值文件，文件以纯文本形式存储表格数据，包括数字和文本。所以想输出成表格只需要按照文本的样式来进行写操作即可。

1. **数据排列**

使用结构体数组来存储各引脚的信息，对照里面的电流值元素，可按冒泡排序法对电流值进行再次排序，每进行一伦筛选可选出最大电流值的引脚结构体信息，保存为另一个结构体数组。多重循环下来即可生成按电流值大小排序的规律信息，可再次输出到csv文档。

1. **系统框图**

测试文本名称识别

进入log文本

识别芯片引脚Leakage数据

**Leakage有效数据的识别**

**建立Excel表格模板**

确定Leakage数据类型

生成表格标题

多文本Leakage数据整合

Leakage数据写入Excel表格

Leakage数据排序

**Leakage数据提取与处理**