**2019.3.4**

第一次文档主要是讨论以下几个内容

ps.因为我很菜，所以需要各位都给予意见并且配合工作任务，争取能最终弄出一个能看的东西

* 需求分析

基本计划: 先把sql到最终的查询计划这一块实现出来，然后研究与下层的交互

(本周还会线下开一次小会讨论以下，并且会和Tikv的人一起讨论)

* 语言的选择

因为目前大家的语言栈都不相同，而结合分布式以及TiDB，候选在C++，Java，Go当中选取一种

C++学习梯度相对大，指针麻烦

Java入手相对快，语法糖多，也算比较主流

Go入手相对快，和TiDB源码契合，但语法糖少，没有类等。

个人推荐Java。。但是各位好像对C++比较熟悉哦⊙∀⊙！

这个还可以商凉。。。

* 资料调研

首先如果不熟悉查询优化，可以大概看一下查询过程，知道物理优化逻辑优化里面包含什么

接下来是SQL部分的调研，主要是把输入一个sql到parser解析这一个过程弄明白是怎么实现的，大家可以一起找一个demo(这个很关键)

最后是之前说的TiDB的CTO黄东旭的公开课，可以参考一下，链接我发到群里

* 分工

明天会在github建仓库，然后各位学习一下怎么进行团队协作，多用终端，少用GUI，具体分工我们后面再讨论

**2019.3.7**

今天主要尬聊的是以下几个问题：

* 关于git团队协作的资料可以参考这两个，非常的详细简单易懂，对于不熟悉的git命令可以参考官方文档或者廖雪峰的博客或者直接在命令行通过man查看。我已经建立好了repo

<https://www.jianshu.com/p/3cd9b85a88ae>

<https://blog.csdn.net/ying422/article/details/45131079>

<https://www.cnblogs.com/schaepher/p/4933873.html>

* 语言确定，这次开发使用Java8进行，需要花一点时间补一下的内容如下

<<阿里巴巴Java开发手册>> 主要统一小组的代码规范

<<Think in Java>> 暂时只需要看基本语法

* 分工说明：目前仍为报团行动
* 下一步计划(3.7-3.14)

1. 继续调研sql parser部分，熟悉回顾编译原理基本知识(BNT等等)

2. 了解Lex & Yacc

3. 尝试把Druid Parser(德鲁伊)跑起来，通过文档了解基本设计思路

4. 继续看TiDB博客，寻找小demo

5. 设定好token，sql映射

**2019.3.14**

Google F1: http://www.10tiao.com/html/421/201810/2247486921/2.html

问题讨论：

1. 词法分析的实现：

(1) Token分类，使用正则

(2) 暴力方法，使用循环分词

2. 语法分析的实现：

demo : https://github.com/antlr/grammars-v4/tree/master/mysql

demo : https://github.com/igarzik/sqlfun

通过对比Druid，Antlr，Lax&Yacc，决定放弃L&Y的LALR形式，使用Druid或者Antlr，

采用LL(\*)，从左到右分析，从左开始推导，自顶向下，递归下降

3. 词法分析的实现：

目前的构想是放在生成语法分析树的过程中，问题是需要有表确定表是否存在，对于建表模拟需要进一步讨论

下一步计划(3.14-3.21)

1. 继续调研sql parser部分，熟悉回顾编译原理基本知识(BNT等等)

2. 学习demo，解决未想通的问题

3. 开始尝试编码

by tuuna