



注意:

1. where part每个where存一个文件，否则存一个文件会所有的where在一个mapper里面执行;
2. n个where对应n个mapper，1个reduce

方案缺点:

1. 调度需要使用专门类似于鸟贼这样的组件;
2. 网络隔离性，希望hadoop环境 and 外网是隔离开的;