|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 个人简历 | | | | | | | | |
| * 基本信息 | | | | | | | | |
| 姓 | 名： | 陈俊伊 | 性 | 别： | 女 | | |  |
| 毕业院校： | | 东北大学 | 学 | 历： | 硕士 | | |
| 毕业年限： | | 2 年半 | 毕业时间： | | 2019.01 | | |
| 年 | 龄： | 29 | 籍 | 贯： | 辽宁省抚顺市 | | |
| E - Mail： | | [381782094@qq.com](mailto:381782094@qq.com) | 联系电话： | | 18842595746 | | |
| 工作年限： | | 5 年 |  | |  | | |
| * 毕业院校 | | | | | | | | |
| 本科 | | 辽宁中医药大学（信息管理与信息技术） | | | 2012.8 | - | 2016.7 | |
| 硕士 | | 东北大学(软件工程) | | | 2016.9 | - | 2019.1 | |
| * 在校期间奖惩情况 | | | | | | | | |
| 本科期间获得过一等、二等奖学金，并在大三下学期期间获得保研资格；  获得 2016-2017 学年东北大学校级奖学金一等奖；  获得 2017-2018 学年东北大学校级奖学金一等奖；  获得 2018-2019 学年东北大学校级奖学金一等奖； | | | | | | | | |
| * **个人技能** | | | | | | | | |
| 1. 主要使用java语言开发，在近一年中开始使用c++进行开发 2. 熟悉JVM内存结构、内存分配与回收策略 3. 熟悉数据结构 4. 熟悉Linux操作和shell脚本 5. 熟悉Elasticsearch、Lucene相关知识，查看并修改过源码 6. 熟悉clickhouse写入和读取逻辑， 阅读过写入和读取源码 | | | | | | | | |
| * 工作经验 | | | | | | | | |
| 公司名称：奇安信科技集团 | | | | | | | | |
| 工作时间：2019 年 2 月末-至今 | | | | | | | | |
| 工作职位: ES 研发运维、lucene代码熟悉、clickhouse代码熟悉和二次开发 | | | | | | | | |
| 工作描述：负责公司内部和客户现场的 ES 集群运维，内部负责维护 ES 版本更新，ES 优化，c语言版全文引擎的开发和运维；clickhouse减少merge次数，进行追加写入的调研和研发工作 | | | | | | | | |
| * **项目经验** | | | | | | | | |
| 一、项目名称：ES集群运维 | | | | | | | | |
| 项目职责：  负责内部集群和外部客户ES集群的运维工作，通过运维工作了解和学习ES的原理 | | | | | | | | |

|  |
| --- |
| 二、Elasticsearch二次开发 |
| 项目职责：根据公司业务需要，对Elasticsearch的源码进行过阅读，并根据业务需要进行过二次开发，分别有：   1. 某海外业务为了打开更多的索引，修改正排索引为只加载7天内的数据，超过时间的数据不加载，现用现加载，并在一个时间段内不再有查询请求进行卸载操作 2. 添加forcemerge索引级别的配置 3. 对于数值类型的查询，在lucene层面添加跳段功能，以减少segment的扫描。提高查询速度 4. 为了提高查询响应，在只有查询没有聚合的查询语句中可以通过添加hits\_partial配置，目的是当有一个或者多个分片查询到满足要求的数据后，直接返回给客户端，不再等待所有分片的数据，来减少查询时间 5. 查询截断功能：Elasticsearch原生就有terminated\_after的功能，但是因为Lucene在查询时要对查询到的文档进行一个相关度的计算，故需要在查询阶段扫描目录下的全部segment文件，在收集阶段执行了真正的截断，在某些不需要打分的场景中，增加了查询的时间，并且增大了IO的压力，针对不需要进行打分，只需要将最近的前几条数据取回的场景，添加了真正的查询截断功能，在查询阶段查询到满足size的要求后，就停止本次查询，不再扫描全部segment文件 6. 延迟索引功能，因为业务的字段大多有索引需求，但又不会都进行实时查询，可能频繁查询的字段只有几个，只有在护网或者特定情况下才会进行全字段的查询需求，根据这些业务的需求，添加了延迟索引功能，即在索引写入时，只将确定频繁查询的几个字段的index:true，其他不查或者很少查询的字段不进行索引，当需要对这些字段进行查询时，通过添加delay\_index配置，进行延时索引，以提高写入速度 |
| 三、借鉴lucene，开发c语言版本的全文引擎库 |
| 项目职责：为了避免java的GC，并适配公司的业务，使用c++将lucene的索引和查询过程进行翻译。主要负责storedField类型字段的索引和查询流程，索引过程中merge选段和执行流程，以及短语查询等流程。 |
| 四、clickhouse追加写入功能调研和研发 |
| 项目职责：clickhouse一个insert操作就会生成一个part，part过多除了会导致在查询阶段需要扫描的目录过多外，还会经常触发merge操作，导致IO资源紧张。为了减少merge操作的次数，调研clickhouse的写入和读取逻辑，在写入阶段添加追加写入功能，即选择一个满足要求的part进行追加写入，避免每次写入都生成新的part，从而降低merge的次数 |
| * 自我评价 |
| 1. 有较强的学习能力，并有学习意识，能主动学习； 2. 有较强的时间和交付观念； 3. 平时喜欢跑步，打羽毛球等运动，乐观积极； 4. 有阅读技术书籍或者博客或者视频的习惯，保持持续学习的态度 |