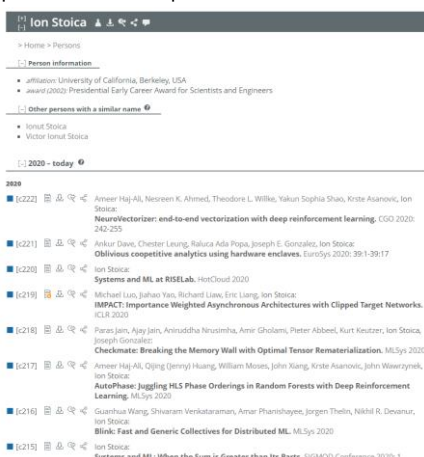
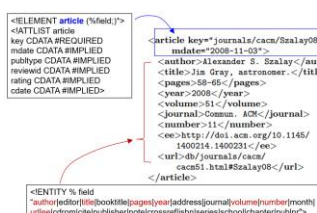


编程作业 2: 分布式 DBLP 数据查询系统

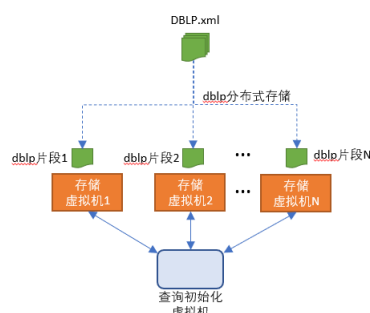
1. **DBLP 数据集介绍.** DBLP (Digital Bibliography & Library Project)是计算机领域内, 其 url 链接为 (<http://dblp.uni-trier.de>), DBLP 研究的成果以作者为核心的一个计算机类英文文献的集成数据库系统, 按年代列出了作者的科研成果, 包括国际期刊, 会议, 专著等公开发表的论文。这个项目是德国特里尔大学的 Michael Ley 开发和维护, 它提供计算机领域科学文献的搜索服务, 但只储存这些文献的相关元数据, 如标题, 作者, 发表日期等, 并没有包括文献的全文文件 (如 pdf 文件)。下图为加州大学伯克利分校教授 Ion Stoica 的 DBLP 论文列表 (对应的 url 为: <https://dblp.uni-trier.de/pid/s/IonStoica.html>)



DBLP 使用 XML 存储元数据 (<http://dblp.uni-trier.de/xml/>), 对应的 dblp.xml.gz 文件大小为 553 MB。DBLP 在学术界声誉很高, 很多论文及实验基于 DBLP 开发。所收录的期刊和会议论文质量较高, 也比较全面, 文献更新速度很快, 能很好地反应了国外学术研究的前沿方向。如下图为一个 XML 文件格式示例 (见 <http://dblp.uni-trier.de/xml/docu/dblp.xml.pdf>)



2. 给定一个查询初始化虚拟机和 N (如 6) 个存储虚拟机, 其中 N 个虚拟机负责分布式存储上述 dblp.xml 片段文件, 查询初始化虚拟机负责发送一个查询请求, 协调 N 个虚拟机, 完成如下查询功能需求:



给出输入条件: 作者名字 author 如 Ion Stoica, 返回作者的 DBLP 发表论文总数。

3. 非功能性需求：
 - a) 存储负载均衡: 每个虚拟机存储 dblp.xml 的部分片段内容, 要求每个虚拟机所存储的内容空间将可能接近, 方法可以按照每个文章条目随机均匀的分配到任意一个虚拟机 (这里我们假设每个文章条目的 xml 内容存储空间是一样的);
 - b) 查询容错机制: 这里假定查询初始化虚拟机不会出现故障、而 N 个存储虚拟机存在故障的可能性, 要求该查询在 N 个存储虚拟机出现故障的情况下还可以正确的查询结果;
 - c) 最小化端到端的查询时间: 在满足上述 a)和 b)两个非功能需求的情况下, 尽可能的降低在提交查询之后和返回最终查询结果之间的查询时间。
4. 提示:
 - a) 通过 socket 方式实现 client/server 网络通信模式, 在每个存储虚拟机实现一个 server 程序, 在查询初始化虚拟机实现一个 client 程序, 初始化查询请求发送至每个 server; server 程序接收到查询请求后, 负责查询本地 dbml 片段文件, 返回查询结果给 client 端, client 端则需要汇聚 N 个查询返回结果, 最终得到汇总后的查询结果。
 - b) dblp.xml 的分布式存储: 简化起见, 有关 dblp 论文的片段分区准则, 可以考虑将 dblp.xml 随机切分, 然后将一个 dblp 分片随机存储在 N 个存储虚拟机之一, 以达到负载均衡目的; 为了保证容错机制, 每个 dblp 分片额外存储另外一份副本, 使得每个 dblp 分片最多有两个副本, 并且这两个副本分别存储在不同的虚拟机。在进行查询处理的时候, 避免由于两个副本的缘故导致查询结果重复的问题, 因此需要对容错副本进行区分。
 - c) 查询时间的优化: 简化起见, 考虑到本次作业仅涉及到查询 author 信息, 只须采用类似于 wc 命令统计输入作者 author 的频次; 可以在每个存储虚拟机统计包含输入 author 的频次, 然后存储虚拟机将本地的 author 频次返回至查询初始化虚拟机, 然后由存储虚拟机进行汇总, 最终得到查询结果。
5. 查询初始化虚拟机最终生成 1 个 log 文件: 学号-hw2-q1.log, 该 log 日志文件最后一行为所用查询的时间。
6. 评分准则 (共计 15 分):
 - a) 查询 Demo 演示: 查询结果正确 (4 分)、存储负载均衡 (3 分)、查询容错 (3 分)、在规定时间内 (如 1 分钟) 内返回结果 (3 分)
 - b) 作业报告: 2 页, 作业报告包括: 组成员算法设计、如何支持可扩展性、低延迟、消息格式、作业 2 如何与本次作业集成、网络带宽开销 (节点 join、leave、failure-rejoin)、消息丢包时组成员列表的错误比例。(2 分)
7. 提交文档包括 3 个内容: 源代码文件、日志文件、作业报告。文档以“学号-hw2”.zip 命名。