**YourSQL: A High-Performance Database System Leveraging In-Storage Computing**

主机处理器＋固态硬盘嵌入式处理器

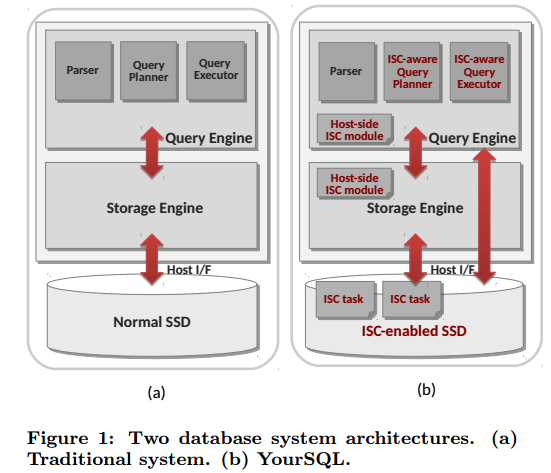
（也有用到专用硬件加速器（即硬件模式匹配器），但主体由嵌入式处理器实现）

应用领域：数据库

IO紧凑型查询希望能够快速的返回最终结果。一个加速查询的方法是减少存储设备和主机系统间传输的数据量，这可以通过1）早期过滤，即过滤掉无关数据或2）仅传输中间和/或最终计算结果来实现。

早前早期过滤通过预设过滤元数据来实现，元数据管理会带来开销，且其实质并未减少存储设备和主机系统间传输的数据量，因为在过滤前，数据均需被传递到位于数据路径（存储设备和主机系统间）上的过滤服务器或FPGA上。

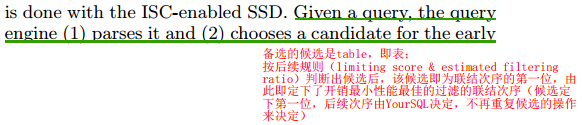
本文提出YourSQL，通过卸载查询的数据扫描操作到SSD来实现早期过滤。YourSQL上的查询是以分布式方式由主机系统和SSD共同处理，经设备端早期过滤后，仅将相关页面传输给主机端进行处理，实现存储设备和主机系统间传输的数据量的减少。为易于实现并考虑到与现存系统的兼容性，YourSQL结构完好保有传统数据库系统结构，如图1所示。

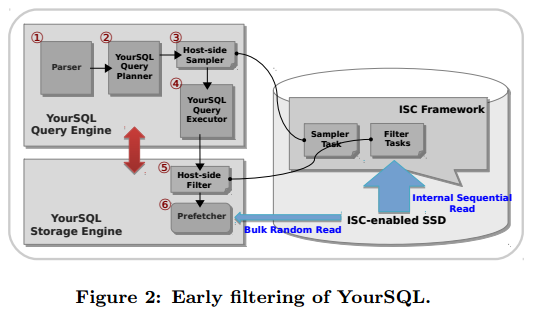
（a）中传统系统由数据库系统（查询引擎和存储引擎）和SSD组成。对于一个查询，查询引擎设置查询计划并执行，存储引擎处理来自查询引擎的IO请求。（b）YourSQL大体上符合（a）的传统结构，但为实现ISC，ISC使能的SSD与整个软件栈是通过ISC模块感应的。

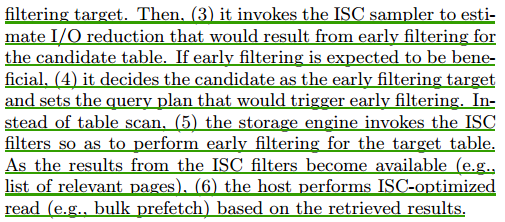
YourSQL是在Biscuit的结构上实现的，为实现ISC，具有如下特点：

1. 用户可编程SSD允许复杂查询操作以ISC任务的形式被卸载。→ ISC任务形式即Biscuit中的SSDlet执行单元，是用户可编程定义的
2. 为激活设备端的ISC任务并接收来自设备端的结果，主机端和设备端之间需要接口。→ Biscuit结构已实现交互接口
3. 为实现ISC进行查询计划优化。→ 这里可通过ISC任务编程实现，原理后续会详细介绍
4. 为ISC进行数据路径的重新组织。 → 原理后续会详细介绍

YourSQL进行早期过滤的总体流程如下：

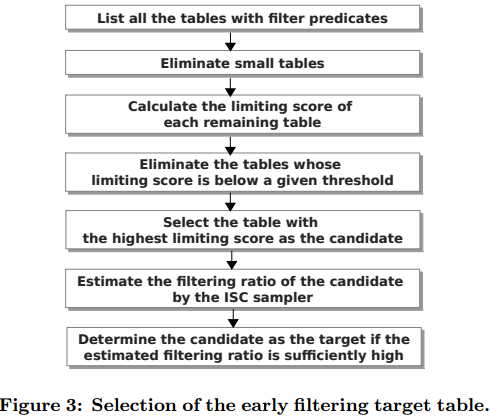


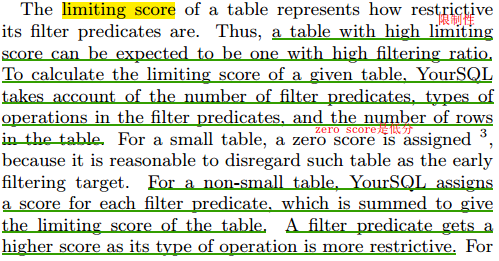




**(1)(2)步：候选目标表的选择**

YourSQL以过滤率（filtering ratio）为选择标准来选择早期过滤的目标表（早期过滤若能带来更高的IO减少，则会更高带来性能的提升←以过滤率为标准，在所有表中挑选出受限制最严格的表作为目标，并仅传输高度相关的页面，提升IO减少量）。过滤率是页面级别的选择度，表示为页面集中（与查询）相关的页面的比率。其值由0到1，其中0为最高过滤率，表示经过滤后无相关页面，即过滤程度最高。过滤率取决于过滤谓词（即条件谓词）。由过滤率定义可以看出，其与查询判断条件高度相关，以其为选择标准，能筛选出高度相关的页面，提高性能。

若简单通过扫描整个表并计数符合查询条件的页面，开销极大。为在合理开销下选择具有高过滤率的目标表，引入limiting score和estimated filtering ratio。YourSQL以limiting score为标准选择候选目标表（步骤（2）），再以estimated filtering ratio为标准确定目标表（步骤（3）（4））。

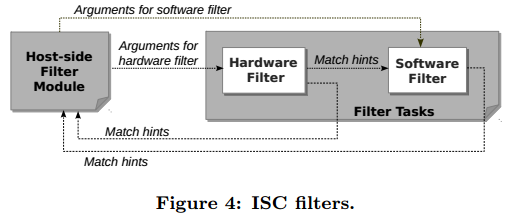


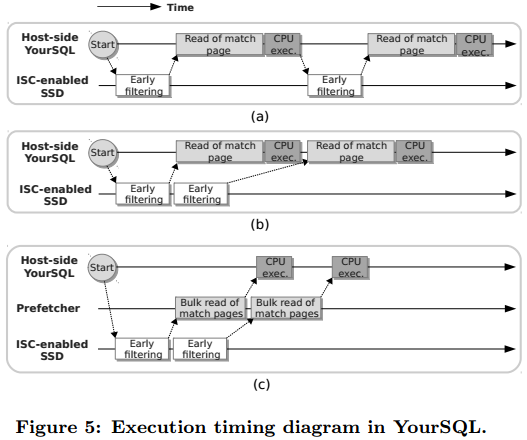
**（3）（4）步：目标表的选择**

ISC sampler评估estimated filtering ratio以确定早期过滤的最终目标表。ISC sampler功能原理与硬件过滤器（Biscuit结构中的硬件模式匹配器）相同，都是扫描，不同的是扫描的范围，ISC sampler仅扫描部分表，estimated filtering ratio则是其中匹配页面的比例。

**（5）步：执行早期过滤**

确定早期过滤的最终目标表后，则需将该表对应的过滤条件卸载，是YourSQL设备端能够根据这些过滤条件进行检索和匹配，这就是filtering condition pushdown(FCP)。FCP经由硬件过滤器（Biscuit结构中的硬件模式匹配器）和软件过滤器的双重过滤（因为硬件模式匹配器的过滤粒度不够小，可能会存在误报），如下图4所示。

早期过滤返回的是一个匹配提示数组（match hints array）。对表的每个页面进行条件过滤匹配，若页面与该条件匹配，则置1，直至最后一页，若目标表对应的过滤条件不止一个则重复上述操作。匹配提示数组返回给存储引擎以避免无关页面的无关访问，进而减少IO。

**（6）步：主机端操作**

YourSQL将查询操作分三步处理：早期过滤，匹配页面读取，行处理。这些步骤可同步执行（如下图（b），下图（a）为顺序执行的情况）。其中匹配页面读取可通过预取操作减少时间开销（如下图（c））。

Summary:

韩国三星公司于2016年在Biscuit结构的基础上实现了一个利用存储内计算技术加速查询的数据库系统原型，具体方式是在一个用户可编程的固态硬盘内同时利用ISC、可编程性及其内置的模式匹配硬件模块实现数据过滤功能，减少硬盘到主机端的数据传输。实验结果表明，相对于传统的数据库系统，该系统能够将查询性能最多提高15倍，将能耗降低近24倍。