

借鉴 Oracle , 深入修改 MySQL/PostgreSQL 内核代码

---- PG篇

【项目背景】

1、掌握研究开源数据库的方法，构筑自主可控的知识体系。

国产数据库百花齐放，数据库生态欣欣向荣，为技术人员提供了历史性的机会。对开源数据库的研究，渐渐成为大势所趋。但成熟的开源数据库，动辄上百万行代码，让企业与技术爱好者无从下手。在学习、研究过程中，企业与技术爱好者大都面临两个重要问题。

1) 超长时间的学习、研究周期，而且在期间没有产出（只有时间、人力的投入）

2) 没有参照、没有对比。无论是源自图灵奖得主斯通布雷克的 PostgreSQL，还是应用广泛的 MySQL，它们的代码一定都是千锤百炼的。我们何德何能，敢对这些高品质代码动刀！

针对第一个问题：学习/研究周期长。可以用“基于调试数据库的正向反馈”法解决。这种方法的介绍，将是本课程的重点。我们不可能在十数小时课程中，将百万行的代码历数一遍。因此，分享阅读源码的方法，才是本课程的重点。“基于调试数据库的正向反馈”法，可以让我们在漫长的学习/研究周期中，不断有成果、有产出。比如，为开源数据库提供像 Oracle 一样测试、观察性能的工具等。这种产出可以奖励我们不断前进下去。

授人以鱼，不如授人以渔。本课程将重点放在“渔”上，为有志于在开源领域发展的企业和技术爱好者，提供方法。

至于第二点，“没有参照”。其实有一个很好的参照：Oracle。若论技术上的先进性，Oracle 无疑仍是最先进的数据库之一。而且 Oracle 累积了大量的成功经验。参照 Oracle 的原理与机制，对 PostgreSQL/MySQL 做一些调整，最终达到自主、可控，将是成功的最佳路径。本课程将以 Oracle 部分原理为例子，实例讲解如何参照这些原理，修改 PostgreSQL 或 MySQL。

2、掌握用户态、内核态动态跟踪的方法，构筑自己的职业壁垒。

对于开发、DBA，或是其他数据库相关从业者，构筑自己的技术壁垒十分重要。

本课程的“调试数据库”方法，将综合运用操作系统原理与数据库原理。这将使我们对原理不仅仅停留在了解之上，而是要进一步的运用。“纸上得来终觉浅，绝知此事要躬行”。只有去运用从“纸”上学到的原理，这些原理和变成自己真正掌握的知识。而且，使用调试数据库方法，在以后的学习、研究过程中，还将不断加强自己对计算机体系底层、

操作系统原理、特别是数据库原理的掌握与运用。这种底层原理性知识，将形成自己的技术壁垒，使自己在 IT 领域立于不败之地。

3、补足基础、夯实地基

数据库毕竟属于系统软件，对计算机体系了解也要求更高，考虑到有部分有兴趣的学员基础不足，本课程将附赠 10 个小时的基础课程，包括“变量是什么”、“堆栈是什么”、“HASH 是什么”、“BTree 是什么”、“开发简易 SQL 解析器”等等，为大家补补基础，提升大家对计算体系的了解程度。

基础课程将以录播形式，方便大便反复观看。部分内容锻炼动手能力，比如，我们要自己动手写一个 BTree、亲自写一个简易 SQL Parse 解析器等等。通过动手，提高对计算机体系的整体了解程度，为后面源码研究打好基础。

4、除了本次深度培训，还会提供后期的跟踪问答

虽然已经相当深入全面了，但如果参加受训的你，真正开始调试数据库，可能还会遇到很多问题，针对这种情况，在 ITPUB 社区，我们会针对这个话题，进行后期的针对性问题的有偿问答服务，让你真正掌握调试数据库技术，可以使用此种技术去分析、甚至是修改一个开源数据库的内核。

【课程介绍】

数据库技术的发展，在国内已经进入下半场。很多企业不再停留在单纯的使用与运维，“自主、可控”已经是所有企业的共识。在需求推动下，对数据库内核技术的研究，也如火如荼的展开。本课程主要为大家分享使用调试技术研究数据库内核的方法，以及如何将研究结果应用到日常场景，以帮助我们更好的运维、使用数据库，满足应用需求。并且重点为大家分享如何参照 Oracle 在超高并发、极短响应时间的成熟经验，对 PostgreSQL/MySQL 的源码在某些方面进行改进。

本课程整体上分为两大部分：

1. 基础部分
2. PG 部分

基础部分重点为大家补足基础。

PG 部分以内存管理、xlog 体系为例，重点讲述这两部分的原理、相关代码，以及如何进行改进。

内存管理部分，在原理的基础上，将进一步讲述如何发现代码中的 BUG，以及如何绕过或开发 patch。

Xlog 体系部分，讲对照 Oracle，重点讲述如何参照 Oracle 的体系，增加 PostgreSQL 的性能。

根据讲师的经验与授课体系，有意报名的参课者最好有如下的一些准备或者经历：

1. 有 Oracle/MySQL/PostgreSQL 至少一种数据库使用经验，并至少对其中某一数据库常见原理有基本了解。
2. 对 Linux 操作系统常见原理有基本了解。
3. 对 Linux 的编程有一定的了解（Linux 下 c 开发最佳），如果没有涉及过，请提前准备相应的资料，并提前准备学习一下。

【课程收益】

1. 掌握计算机基础知识，理解程序、内存、指针、堆栈等。
2. 理解什么是数据结构，掌握链表、HASH 等基础的数据结构。为学习开源数据库源码打下基础。
3. 了解 SQL 的解析过程，通过亲手构造一个简易版 SQL 解析器，消除对数据库体系复杂度的恐惧心理。
4. 了解 CPU 层的优化常识，为发掘、提升开源数据库代码打下基础。
5. 掌握调试数据库技术。
6. 掌握阅读开源数据库代码的方式。
7. 掌握分析数据库原理的方法。
8. 掌握使用动态跟踪技术为 PostgreSQL 增加性能观察工具的方法。
9. 掌握影响高并发数据库的要点，以及如何衡量一个数据库的优劣。

【课程大纲】

基础知识部分：(基础知识部分为免费学习部分，课程会录制好放在 ITPUB 学院)

时间 (每课时 60 分钟)	项目名称	分类提纲	内容介绍
学前预备	环境要求	Linux gcc gdb	准备环境
0.5 课时	开篇语	开篇语	课程体系介绍与学习方法介绍
0.5 课时	基础部分	gdb 使用小结	掌握计算机基础知识，理解程序、内存、指针、堆栈等。为进一步学习打下基础。
0.5 课时		程序是什么	
0.5 课时		变量是什么	

0.5 课时		内存是什么	
0.5 课时		堆栈是什么	
0.5 课时		数组是什么	
0.5 课时		堆是什么	
0.5 课时	数据结构部分	链表是什么	理解什么是数据结构，掌握链表、HASH 等基础的数据结构。为学习开源数据库源码打下基础。
0.5 课时		HASH 是什么	
0.5 课时		BTree 是什么	
0.5 课时		自旋锁是什么	
0.5 课时	SQL 解析器部分	Flex 介绍	了解 SQL 的解析过程，通过亲手构造一个简易版 SQL 解析器，消除对数据库体系复杂度的恐惧心理。
0.5 课时		Bison 介绍	
0.5 课时		练手：多项式计算器	
0.5 课时		了解 SQL 的编译：简易版 SQL 编译器	
0.5 课时		PG 的 SQL Parse	
0.5 课时	优化部分	perf event 接口	了解 CPU 层的优化常识，为发掘、提升开源数据库代码打下基础。
0.5 课时		CPU 指令与周期统计	
0.5 课时		CPU 性能指标统计：IPC，每周期指令数	
0.5 课时		L1 cache/分枝预测命中率	

PostgreSQL 部分

时间 (每课时 60 分钟)	项目名称	分类提纲	内容介绍
学前预备	源码环境配置	PostgreSQL 源码编译 cscope、vim 环境配置 System tap 安装	参照讲师文档、视频资料，配置学习环境。 此部分内容，不在课程时间之内。
3 课时	PG 内存分配的 BUG 与 Patch	PG 与 Oracle 内存模型 Freelist 与内存分配/回收 内存上下文 AllocSetAlloc 分析 Freelist 中的 BUG 分析 Patch 方案	1、学习 PG 与 Oracle 内存管理 2、PG 内存管理源码介绍 3、PG 内存管理中的 BUG 分析以及 Patch 方式
3 课时	PostgreSQL 逻辑读研究	PG xlog 与 Oracle 的生成流程 XLogInsert 函数分析 LSN 的变化轨迹分析 PG xlog 日志格式分析方法	PG xlog 的原理与 Oracle Redo 的原理介绍
3 课时	检查点与刷新日志	Checkpoint 的目的与意义 Checkpoint 工作流程 XLogFlush 写日志开始 XLogWrite 写日志步骤分析	1、数据库检查点的意义，PG/Oracle 检查点的原理分析。 2、PG、Oracle 刷新日志缓存的原理与 PG 相关的源码介绍
3 课时	PostgreSQL 解析过程研究	Full Page Writes 作用分析 Full Page Writes 关闭时的 xlog 处理流程 20%到 50%性能损耗：Full Page Writes 的影响 参照 Oracle 的改进：数据高可用与块恢复	参照 Oracle 体系，以最少的工作量改造 PG，大幅提升性能实践。

【课程对象】

开发、DBA、系统架构师，对开源数据库技术感兴趣的人员

【授课方式】

讲解结合实操，学员自备可以上互联网的笔记本电脑，所有授课和实践所需的软件、工具、源代码均通过云端服务器访问。

【讲师介绍】



吕海波（花名：VAGE）
美创科技，ITPUB 管理版版主

IT 老兵，25 年 IT 领域从业经历，十数年数据库经验，惯看 IT 江湖风起云涌。曾在多家巨头型互联网公司（阿里巴巴、京东、ebay）从事数据库管理与研究工作。

曾多次在 DTCC 数据库大会、SACC 系统架构师大会以个人、独立身份发表演讲，演讲内容涉及分布式数据库、云数据库与源码研究等方面内容。

在 2019 DTCC 达梦 8 数据库发布会中，作为国内企业级杰出数据库专家代表，与中国工程院院士倪光南、方滨兴共话中国数据库技术的自主可控话题。

出版技术书籍《Oracle 内核技术揭秘》，被誉为国内最深度解密 Oracle 算法原理的技术书籍。现就职于美创科技，从事数据库内核与云数据库、数据安全方面的研究工作。

讲师资质：多届数据库大会演讲嘉宾

著作书籍：

出版技术书籍《Oracle 内核技术揭秘》

正在著作《数据处理技术之旅》（暂定名）