

世界最强大的开源关系型数据库

世界最强大的开源关系型数据库

北京大学PostgreSQL内核开发入门实践V4.2

世界最强大的开源关系型数据库

PostgreSQL中国生态发展战略白皮书

目录

[项目概要 1](#_Toc2871)

[背景意义 1](#_Toc23209)

[企业导师 2](#_Toc9541)

[实践安排 3](#_Toc22597)

[课程计划 3](#_Toc3511)

[课程资源 6](#_Toc11584)

[预习内容 6](#_Toc19838)

# PostgreSQL数据加密与存取开源开发实践

# 项目概要

**主要人员：**

瀚高研发人员2人、社区技术专家1人、北大学生小组预估6人以上；

**项目周期：**

16周（时间自主，建议每周固定时间，每周至少保证2小时）

**开展形式：**

企业资深研发人员与学生以线上协作为主，推进学习计划及作业打分，鼓励学员参与透明数据库加密项目和参与PostgreSQL 国际社区贡献

# 背景意义

**本项目定位为PostgreSQL国际社区重要开发项目。**

以下课程内容主要分为**PostgreSQL数据加密和PostgreSQL数据存取**两个方向，**PostgreSQL数据加密**包括**透明数据库加密（TDE)、密钥管理系统（KMS)**两个项目主题，**PostgreSQL数据存取包括nonFPW实现partial write恢复功能**一个项目主题。**须要大家面向两个方向跟课学习，然后从三个项目主题中选择一个作为大作业，最终通过系统编码理解、应用所学知识。**

**方向一：PostgreSQ数据加密**

PostgreSQL国际社区一直在进行讨论是否以及如何在Postgres中**实现透明数据加密（TDE）**。 许多其他关系型数据库都支持TDE，并且某些安全标准也要求它。

TDE最主要的功能就是用来保护存储在磁盘上的数据来避免恶意人员直接读取或窃取数据库文件，或是偷走整个磁盘导致用户信息遭窃。

目前PostgreSQL支持许多安全的级别来保证数据安全，比如支持TLS网络链接加密来保证客户端和PostgreSQL 服务器间的网络安全，和强大的用户验证功能来保证数据库用户的真实性，但是PostgreSQL缺乏服务器和磁盘间存储的安全加密，这也是为什么TDE目前在PostgreSQL国际社区里被重视的原因。

**方向二：PostgreSQ数据存取**

**数据库I/O读写**最基本单元是page，page大小通常是8K、16K等数值。但OS中page大小是4K。数据库的page大小是OS page的两倍或更多倍，这带来一个经典的partial write问题（有些资料也称为页断裂）。各种数据库为了解决partial write各出奇招。比如，MySQL使用了double write特性（双写）。PostgreSQL使用了Full Pages Write（简称FPW，整页写）。

PostgreSQL FPW特性能很好的解决partial write问题，但是FPW同时带来了巨大的性能负担，使PostgreSQL性能下降至少30%。在写较多且I/O是瓶颈的环境中，FPW甚至可以使性能下降200%。MySQL的double write也有同样的问题，目前主流保证数据一致性的、完整性的数据库中，只有Oracle较好的解决了这一问题，既避免partial write又对性能没有影响。

PostgreSQL Non-FPW就是针对这一情况，使用类似Oracle的方式，在百分百保证数据一致性、完整性的条件下，既去除了PostgreSQL的FPW特性，又解决了partial write问题，使数据库性能大幅提升。

# 企业导师

**David Zhang**

联系方式：david.zhang@highgo.ca

|  |  |
| --- | --- |
| A person in a suit  Description automatically generated with medium confidence | David是瀚高软件北美研究院（加拿大）的资深软件架构师，在加入瀚高之前，他在智能电网，计量领域，网络安全，资安方面开发创新软件解决方案已有20年以上的行业经验。David于加拿大滑铁卢大学（UW) 获得电气与计算机工程硕士学位，并在以下技术方面拥有丰富的实践经验：网络安全，数据安全，身份验证，软件设计与开发，PostgreSQL数据库功能及内核开发，嵌入式系统，功能和体系结构设计等 |

**Cary Huang**

联系方式：cary.huang@highgo.ca

|  |  |
| --- | --- |
| A picture containing person, person, posing, suit  Description automatically generated | Cary是瀚高软件北美研究院（加拿大）的高级软件开发人员，在加入瀚高之前，他在智能电网和计量领域以C / C ++开发创新软件解决方案已有8年以上的行业经验。 他于2012年在加拿大温哥华的英属哥伦比亚大学（又译“不列颠哥伦比亚大学”，UBC）获得电气工程学士学位，并在以下技术方面拥有丰富的实践经验：高级网络，网络和数据安全性，智能计量创新，Docker部署管理，软件工程生命周期，拓展，身份验证，加密，PostgreSQL和非关系数据库，Web服务，防火墙，嵌入式系统，RTOS，ARM，PKI，Cisco设备，功能和体系结构设计等。 |

**吕海波**

联系方式：lvhb@mchz.com.cn

|  |  |
| --- | --- |
|  | 杭州美创科技数据库内核研究员，"IT老兵，24年IT软件领域从业经历，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、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、云数据库与内核研究等方面内容。 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内核技术揭密》，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。" |

# 实践安排

**PostgreSQL数据加密**

1. 讲解PostgreSQL数据库系统架构、开发环境及开发流程、编译及调试、源代码管理。
2. 讲解加密和算法的基础知识和基本的安全概念。
3. 介绍目前PostgreSQL支持的安全功能与实践。
4. 学习当前PostgreSQL国际社区对TDE开发的内容及流程。
5. 与PostgreSQL国际社区核心成员进行在线交流，讨论国际社区TDE方面的进展
6. 了解当前PostgreSQL对数据的处理，存储及调用的工作原理代码。
7. 在现有的PostgreSQL 的基础上，设计，研究并开发TDE功能以及代码。
8. 学习如何把自己的工作成果分享给PostgreSQL国际社区。

**PostgreSQL数据存取**

1. 讲解PostgreSQL数据库系统架构、开发环境及开发流程、编译及调试、源代码管理。
2. 讲解PostgreSQL数据库I/O与文件体系。
3. 讲解PostgreSQL数据库日志与恢复体系。
4. 了解主流数据库（Oracle、MySQL）的日志与恢复体系。
5. 讲解PostgreSQL数据库xlog生成流程以及相关代码。
6. 讲解PostgreSQL数据库Full Pages Write特性以及相关代码。
7. 讲解PostgreSQL数据库启动流程与恢复相关代码。
8. 讲解PostgreSQL数据库内存管理与相关代码。
9. 讲解PostgreSQL数据库HASH操作相关代码。
10. 在现有的PostgreSQL 的基础上，设计，研究并开发non-FPW以提升性能。

# 课程计划

**以下为PostgreSQL数据加密课时内容，讲师瀚高基础软件研发工程师。**

第一周：

* 课程介绍
* 关于PostgreSQL
* 透明数据库加密和密钥管理
* 数据库常用术语
* 体系架构
* 源码编译
* 使用PostgreSQL
* 内核主要模块
* GDB调试工具
* 概括
* 作业一：了解PostgreSQL(PostgreSQL)、透明数据库加密（TDE)、密钥管理系统（KMS)
* 作业二：搭建你自己的PostgreSQL开发环境

第二周：

* 什么是安全
* 加密学
* 安全的原则
* PostgreSQL的透明数据加密
* 数据在PostgreSQL中的存储
* 访问缓冲区管理器
* 锁定策略

第三周：

* 加密(公钥)
* 什么是密钥管理系统
* 密钥生成
* 密钥存储
* 密钥更新
* 密钥转换
* 密钥撤销和销毁

第四周：

* 介绍
* PostgreSQL源码结构
* 在PostgreSQL中添加一个新模块
* Internal KMS 注意事项

第五周：

* 特别嘉宾：Bruce Momjian，PostgreSQL社区核心成员
* 问答环节
* KMS项目简介

第六周：

* TDE 项目开发实践
* 突发新闻：2021年5月7日，Colonial Pipeline遭遇勒索软件网络攻击
* 提前写日志(WAL)
* TDE的影响
* TDE项目简介
* 小组作业

第七周：

* 软件测试
* PostgreSQL回归测试框架

第八周：

* 什么是PostgreSQL扩展
* PostgreSQL扩展框架
* 创建自己的扩展
* 测试您的扩展

第九周：

* PostgreSQL 社区贡献
* 准备工作
* 补丁审查
* 补丁提交
* 博客贡献
* 总结

第十周：

* 课程总结
* 学生演示TDE项目成果
* 并提交PPT报告
* 设计细节讲解及答疑

**以下为PostgreSQL数据存取课时内容，讲师杭州美创吕海波。**

第十一周：

* PostgreSQL文件体系
* PostgreSQL I/O体系
* PostgreSQL I/O部分源码结构
* PostgreSQL的xlog体系（一）
* Full Pages Write特性（一）

作业：

* 根据源码，总结PostgreSQL 文件与I/O相关数据结构与关系图。
* 使用测试工具，测试、总结Full Pages Write对性能的巨大影响

第十二周：

* Oracle、MySQL数据保护体系
* 了解MySQL Double Write特性
* Oracle为什么不使用Double Write
* PostgreSQL的xlog体系（二）
* PostgreSQL恢复体系
* Full Pages Write特性（二）：相关源码分析

作业：

* 总结xlog生成流程，画出相关数据结构
* 总结Full Pages Write的触发条件
* 总结为什么Full Pages Write会影响性能

第十三周：

* 跟着gdb浏览StartupXLOG()代码
* 恢复前准备工作重点分析
* 恢复相关源码分析
* 恢复后检查工作分析

作业：

* 完成一次备份与恢复操作
* 总结数据库的备份与恢复机制

第十四周：

* 移植StartupXLOG()为nonFPW()
* gdb定位内存上下文错误
* 内存上下文管理
* nonFPW()运行环境构建

作业：

* 自己动手，移植StartupXLOG()为nonFPW()

第十五周：

* 编译按页恢复功能
* PostgreSQL HASH体系介绍
* create\_hash()源码分析
* HASH函数的操作
* 为nonFPW()增加HASH过滤

作业：

* 为nonFPW()实现块恢复功能

第十六周：

* PostgreSQL启动与恢复源码总结
* 主流数据库在partial write的发展
* 主流数据库I/O子系统发展
* Non-FPW未来发展与规化

大作业：

* 为nonFPW()实现partial write恢复功能

**评分标准（考核方式，评分细则）**

* 出勤：10%
* 课后作业：50%
* 大作业：40%

**以下至少选择一个大作业代码实现**

* **大作业1：**TDE代码实现：40%
* **大作业2：**KMS代码实现：40%
* **大作业3：**Non-FPW代码实现：40%

# 预习内容

1. Linux基础

Ubuntu Linux 18.04：https://releases.ubuntu.com/18.04/

Vmware player：

https://www.vmware.com/ca/products/workstation-player/workstation-player-evaluation.html

Virtualbox：https://www.virtualbox.org/

Ubuntu官方Linux指令指南：

<https://ubuntu.com/tutorials/command-line-for-beginners#6-a-bit-of-plumbing>

1. PostgreSQL预习

PostgreSQL github官方代码下载：<https://github.com/postgres/postgres>

PostgreSQL 配置：<https://www.postgresql.org/docs/current/runtime-config.html>

[https://www.postgresql.org/docs/current/auth-PostgreSQL-hba-conf.html）](https://www.postgresql.org/docs/current/auth-pg-hba-conf.html）)

基本PostgreSQL 指南：<https://www.postgresqltutorial.com/>

PostgreSQL内部架构：<http://www.interdb.jp/PostgreSQL/>

PostgreSQL 官方 TDE 和 KMS wiki page：

<https://wiki.postgresql.org/wiki/Transparent_Data_Encryption>

1. SQL语言常用操作
2. GDB使用
3. 主流数据库I/O与partital write

https://www.percona.com/blog/2006/08/04/innodb-double-write/

《Oracle内核技术揭密》第三章：Buffer Cache内部原理与I/O

《Oracle内核技术揭密》第五章：Redo调优与备份恢复原理

# 资源站点

1. PostgreSQL国际社区网站： https://www.postgresql.org/
2. PostgreSQL分会问答网站：[https://www.PostgreSQLfans.cn](https://www.pgfans.cn)
3. PostgreSQL分会资源网站：<https://www.postgreshub.cn>
4. 瀚高软件官方网站 : https://www.highgo.com/
5. 瀚高软件北美研究院网站：<https://www.highgo.ca/>