PostgreSQL 中国生态发展战略

世界最强大的开源关系型数据库

PostgreSQL

CHINA POSTGRESQL ASSOCIATION

目录

瀚高软件数据库加密和密钥管理开源开发实践

项目概要

主要人员：

瀚高研发人员 2 人、北大学生小组预估 6 人以上；

项目周期：

12 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，通过透明数据库加密和密钥管理项目开发实践，鼓励学员参与 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。

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 Transparent Data Encryption（TDE）以及密钥管理 Key Management System（KMS） 功能。 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。KMS 最主要的功能就是管理加密密钥的生命周期，并做定期的密钥更新及撤销。

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，也缺乏了对加密密钥的管理。这也是为什么 TDE 和 KMS 目前在 PostgreSQL 国际社区里被重视的原因。

此课程涵盖密钥管理和数据加密两个主题，学生可以 **选择其中一个主题** 来做设计以及实现。

企业导师介绍

主讲师：David Zhang

	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

主讲师：Cary Huang

	Cary 是瀚高软件北美研究院（加拿大）的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。

特邀导师：Bruce Momjian

	Bruce Momjian 是一位受人尊敬的认真、有趣的长者，他是 PostgreSQL 全球开发小组的联合创始人和核心团队成员，并作为演讲嘉宾参加过诸多国际开源会议。他也是 Addison-Wesley 出版的《PostgreSQL 的简介和概念》一书的作者。 Bruce Momjian 从 2006 年开始受聘于 EnterpriseDB，目前担任 EnterpriseDB 副总裁一职。

实践安排

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 国际社区对密钥管理和透明数据加密开发的内容及流程。
5. 与 PostgreSQL 国际社区核心成员进行在线交流，讨论国际社区密钥管理和透明数据加密方面的进展
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，研究并开发密钥管理或透明数据加密功能。
8. 学习如何把自己的工作成果分享给 PostgreSQL 国际社区。

课程计划

第一周:

- PostgreSQL 数据库基本架构介绍
- PostgreSQL 开发环境搭建，讲解开发流程和要求。
- PostgreSQL 培训代码的编译，调试和问题定位等。
- 使用 Git 和 GitHub 源代码管里工具。
- 培训 GDB 调试工具安装并学习如何使用其来了解 PostgreSQL 的内部逻辑
- 作业一

第二周:

- 数据加密的理论及基础
- 数据加密的数学和概率理论基础
- 传统数据加密的做法和实践
- 数据加密的重要性

第三周:

- 安全级别背景知识基础
- 常用的数据加密算法和应用场景
- 作业二

第四周:

- 密钥管理系统概要
- 密钥管理理论及基础
- 传统密钥管理的做法和实践
- 密钥管理的重要性

第五周:

- 内部和外部密钥管理系统介绍

- 常用的密钥管理和应用场景
- 作业三

第六周:

- 当前 PostgreSQL 加密的发展和使用状况
- 当前 PostgreSQL 密钥管理的发展和讨论
- 常用的加密 C 代码库
- PostgreSQL 数据加密和密钥管理实践

第七周:

- 介绍当前国际社区数据加密的开发情况
- 介绍当前国际社区密钥管理的开发情况
- 观看国际社区核心成员 Bruce Momjian 过去的数据加密和密钥管理开会情况并作解释
- 讲解社区对数据加密和密钥管理的需求
- 社区的数据加密和密钥管理维基百科介绍，资料参考
- 邀请 Bruce 本人与学生进行在线交流，并介绍当前社区情况
- 社区编码标准
- 布置数据加密和密钥管理设计作业（**两者选一**）

第八周:

- 根据学生的方案，给与相应的方向性，和技术性指导
- **选项一：数据加密**
  - 按照项目分组，学生讲解自己的 TDE 开发方案
- **选项二：密钥管理**
  - 按照项目分组，学生讲解自己的 KMS 开发方案

第九周:

- 按照学生的设计方案，提供技术性指导
- 按照学生的设计方案，回答问题，注意事项等
- PostgreSQL 的插件架构介绍与实践

第十周:

- 按照学生的设计方案，提供技术性指导
- 按照学生的设计方案，回答问题，注意事项等
- PostgreSQL 的基本测试框架介绍与实践

第十一周:

- 介绍国际 PostgreSQL 的工作流程和社区贡献指南
- 讲解提交补丁到国际社区的流程（可以考虑向国际社区提交 TDE/KMS 提议或是补丁）
- 列出一些使用 PostgreSQL 的知名企业，还有他们的重点，列举成功案例
- 辅导学生进行社区贡献

第十二周:

- 学生演示数据加密和密钥管理项目成果
- 学生提交 PPT 报告
- 设计细节讲解及答疑

评分标准（考核方式，评分细则）

- 作业：30%
- 出勤：10%
- 密钥管理或透明数据加密设计：20%
- 密钥管理或透明数据加密代码实现：40%

课程资源

1. 可参考得文件，博客及文档
2. 社区设计，讨论，开会纪要的访问
3. 调试工具
4. 可灵活安排线上交流、培训或协同开发
5. 与 PostgreSQL 核心组成员以及关键代码贡献人员进行面对面交流与学习