北京大学 PEKING UNIVERSITY

瀚高 HIGH GO

透明数据库加密和密钥管理

- PostgreSQL开源开发实践

David & Cary

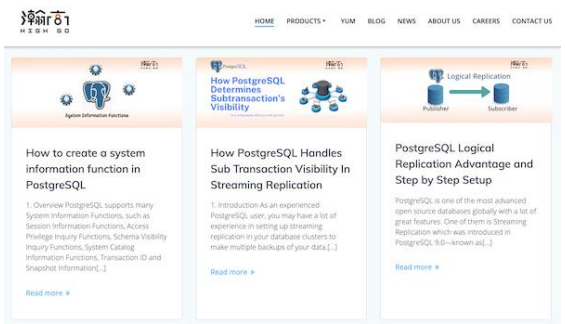
CONTENT
目 录

- 课程介绍
- 关于PostgreSQL
- 透明数据库加密和密钥管理

课程介绍

关于瀚高软件

- 欢迎各位参与这次由瀚高软件提供的开源开发实践课程
- 瀚高基础软件股份有限公司是国内数据库行业龙头企业
  - 国内数据库行业标准主导企业
  - 中国PostgreSQL的主要代码贡献者之一
  - 致力于为政企客户提供有核心竞争力的数据库解决方案
  - 公司始终坚持安全可控和开放创新并重，加强国际交流合作，在数字经济时代为客户的成功提供强大源动力。
- 瀚高软件官方网站 （ https://www.highgo.com/）
- 瀚高软件北美研究院网站 （ https://www.highgo.ca/）



课程介绍

时间安排

- Zoom课程频道
  主题：北大开源PG研发实践课
  时间：2021年3月14日 10:00 上午
  北京时间，每周的星期天，共计12课时
  课程链接：https://zoom.us/j/98626894834?pwd=WU9sdnhmNFhKa2VDZGlYUzNjZU5jQT09
  会议号：986 2689 4834
  密码：HIGHGO

- Teams消息频道

  提供课件，作业，即时消息等

- 课程视频

课程介绍

评分标准

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

课程介绍

我可以学到什么？

- 这是一门围绕PostgreSQL开源数据库的实践课程。
- 本课程侧重于讲解PostgreSQL内部的功能实现逻辑。
- 掌握C语言在Linux平台上的基本编程技能和代码调试工具。
- 学习数据加密和密钥管理的知识基础，应用场景和实践。
- 并把这两种理论知识应用到当前的PostgreSQL开源数据库。
- 了解当前PostgreSQL国际社区的研发及工作模式。
- 与PostgreSQL国际社区核心成员（Bruce Momjian）进行在线交流。
- 个人的课程项目有机会提交给PostgreSQL国际社区成为社区贡献。
- 以后在数据库领域发展和自我的修炼！

课程介绍

议程

- **第一周**：PostgreSQL基础及课程介绍
- **第二周**：PostgreSQL体系架构，编译源码，调试，应用
- **第三周**：数据加密的理论及基础
- **第四周**：数据加密的方法及实践
- **第五周**：密钥管理的理论及基础
- **第六周**：密钥管理的方法及实践
- **第七周**：邀请Bruce本人进行在线交流
- **第八周**：数据加密或密钥管理设计，技术性指导
- **第九周**：PostgreSQL插件架构介绍与实践
- **第十周**：PostgreSQL测试框架介绍与实践
- **第十一周**： PostgreSQL的工作流程和社区贡献指南
- **第十二周**：学生演示数据加密或密钥管理项目成果

导师介绍

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

联系方式：david.zhang@highgo.ca

导师介绍

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。

联系方式：cary.huang@highgo.ca

关于PostgreSQL

关于PostgreSQL

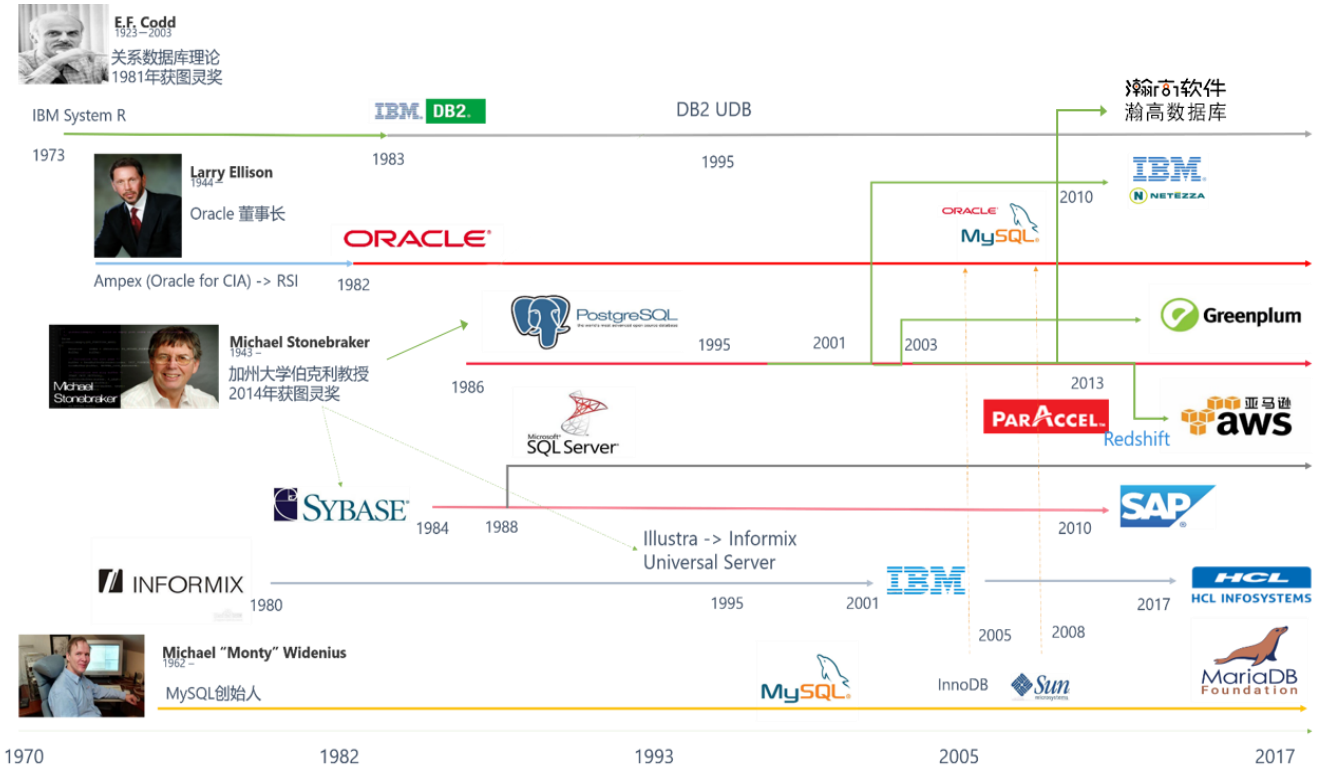
开源数据库介绍

- PostgreSQL是一个功能强大的开源关系数据库系统（Rrelational Database），经过30多年的积极开发，在可靠性，功能健壮性和性能方面都在国际上赢得了极高的声誉。
- PostgreSQL支持多种编程语言，因此可用于跨多平台开发和运行动态应用程序。
- 它被广泛应用于各行各业的顶级公司，例如IT，人力资源，医疗保健，媒体，酒店，教育，电信，金融服务，计算机软件/硬件，广告和营销。
- 它是免费的开源工具，因此在中小企业中同样受欢迎。PostgreSQL的需求现在比以往任何时候都多。
- 很多大型企业都纷纷将它们的数据库系统迁移到PostgreSQL。
- 将PostgreSQL作为一项技能添加到您的个人简历中肯定会帮助您攀登成功的阶梯。

关于PostgreSQL

起源发展

PostgreSQL的起源可以追溯到始于 1977 年伯克利大学的 Ingres 项目，这个项目是 Michael Stonebraker 教授领导。



关于PostgreSQL

Michael StoneBraker

Michael Stonebraker 教授是数据库发展史上几个重量级人物之一，1992年提出对象关系数据库模型，发明了许多几乎应用在所有现代数据库系统中的概念。

2014年，Michael Stonebraker 获得图灵奖（因Google的资助，从本次颁奖起图灵奖金额变为100万美元）。

关于PostgreSQL

国际社区大咖 Bruce Momjian

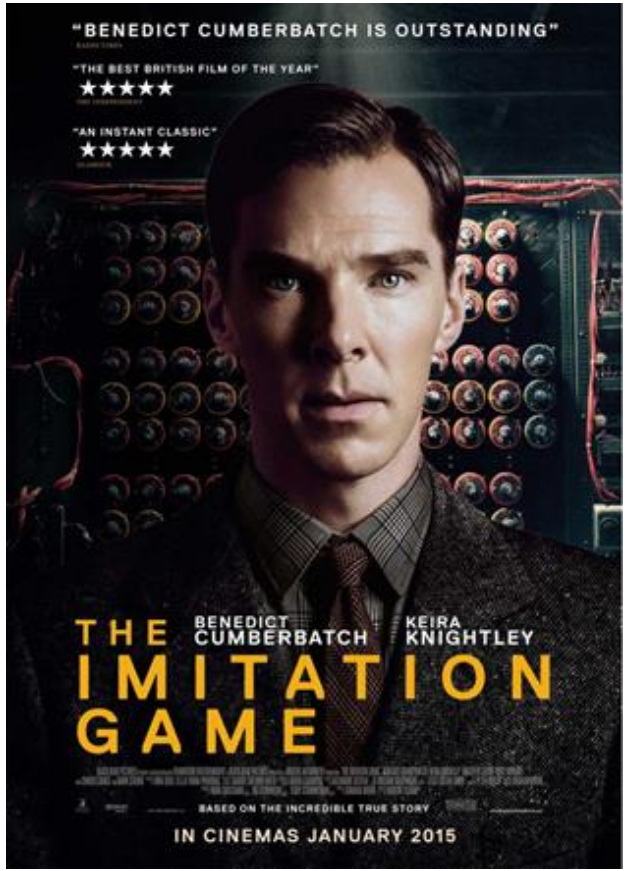
另外一个重量级人物Bruce Momjian是Postgres国际社区创始人，同时也是EnterpriseDB公司的高级数据库设计师。

在之后的议程里，我们会邀请这位国际大咖来和我们交流

关于PostgreSQL

图灵奖-Turing Award

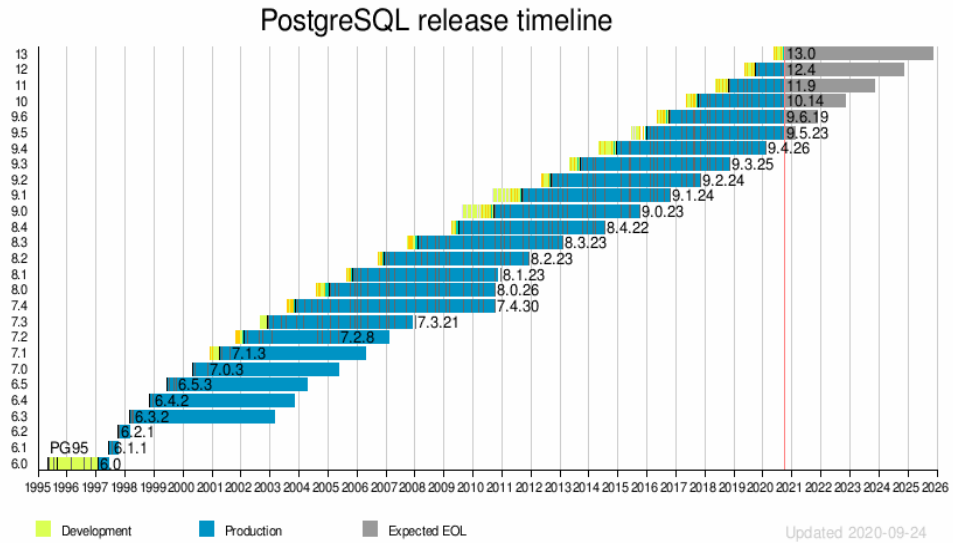
- 图灵奖（ACM A.M. Turing Award），又译杜林奖，是计算机协会（ACM）于1966年设立的奖项，专门奖励对计算机事业作出重要贡献的个人
- 该奖项以艾伦·图灵命名，
- 图灵是一位英国的数学家，在二战时期创造了一台机器来破解纳粹德国用 Enigma machine 加密过后的机密文件
- 2015 年发布的一个电影模仿游戏，就是在说图灵怎么破解纳粹德国的加密文件导致最终德国战败的故事
- 值得一看！



关于PostgreSQL

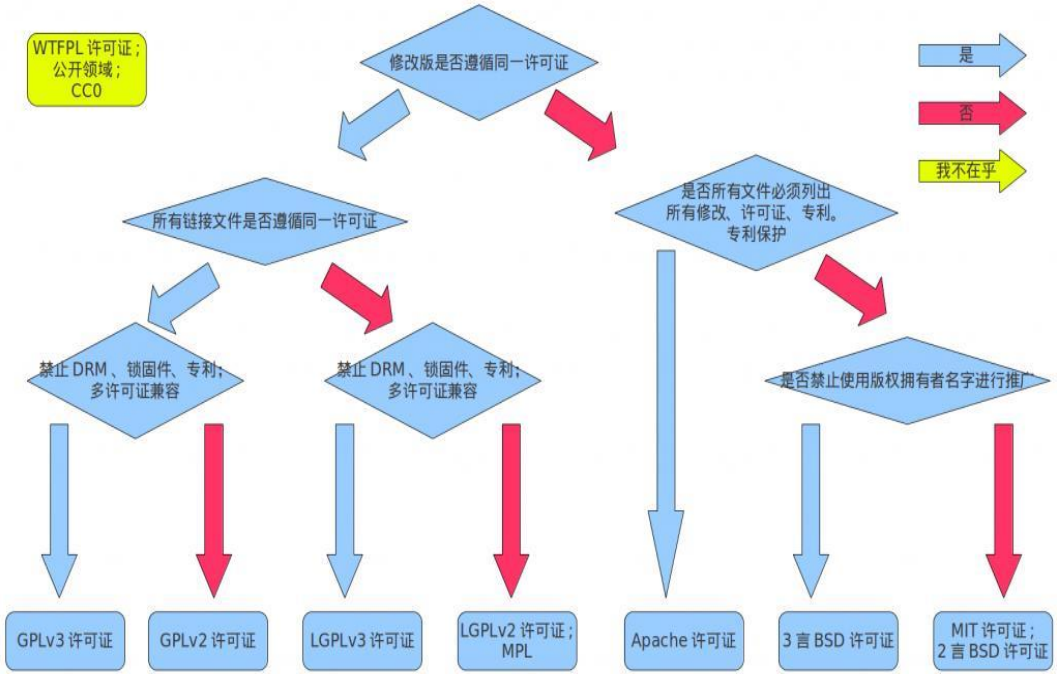
稳定的版本发布

- PostgreSQL是全球最强大的企业级开源数据库
  - 具有既有SQL通用性以及NOSQL扩展性
  - 对芯片友好
  - 版本迭代稳定
  - 国际社区支撑强大
- PostgreSQL全球开发小组于2020年10月24日宣布，PostgreSQL 13的正式通用版本现已可供下载。
- PostgreSQL全球开发人员有500多位，分散在世界各地，彰显了社区强大、稳健的发展节奏和自由、活跃的社区氛围。



关于PostgreSQL

开源协议

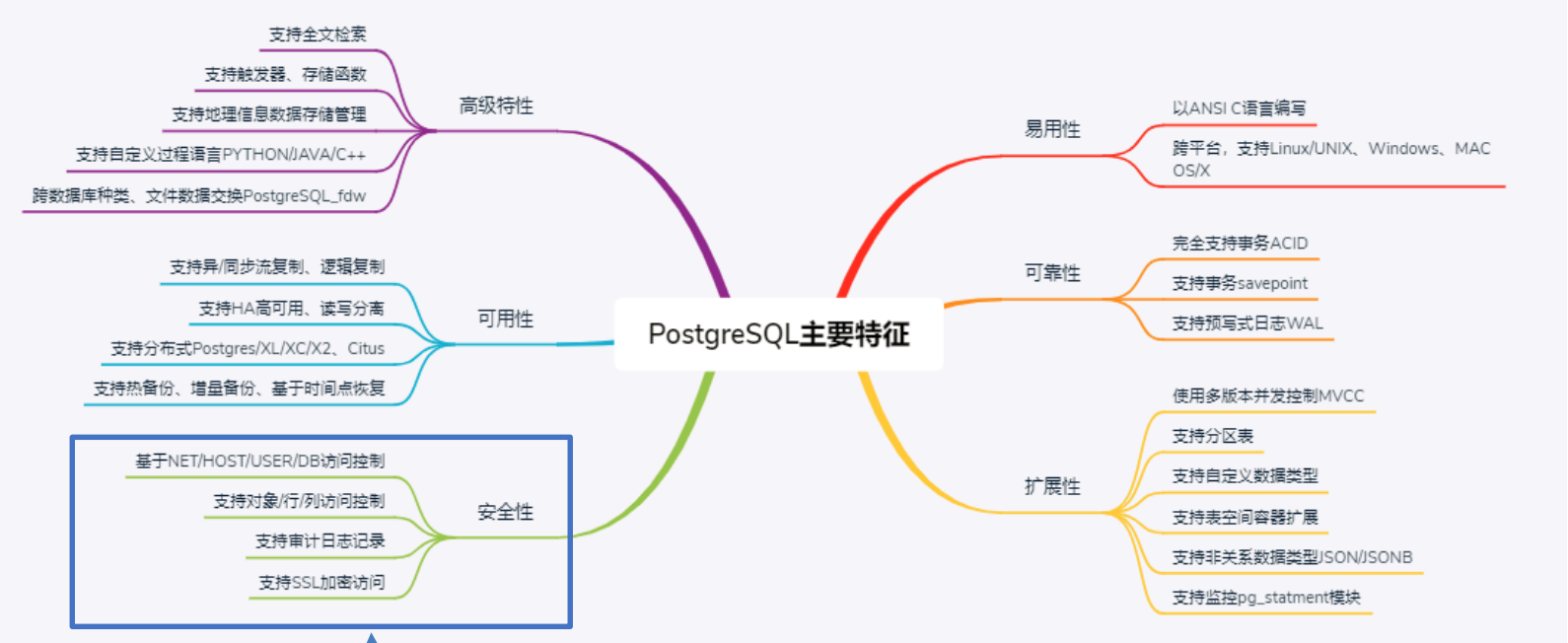


- PostgreSQL的开源协议是类BSD协议或MIT协议，可以在任何目的下进行分发、闭源或者开源。只需将PostgreSQL协议内容附属到源码中即可。
- 开放友善的开源协议，工业级的代码开发，造就了现如今丰富的生态支撑。
- PostgreSQL是中国政、企客户安全合规，开源替代的最佳选择，最佳途径。

原文链接(https://www.postgresql.org/about/licence/)

关于PostgreSQL

技术特性



我们的课程着重
在‘安全性’的话题

关于PostgreSQL

技术特性

- PostgreSQL是一种几乎可以运行在各种平台上的免费的开放源码的对象关系数据库管理系统
- 拥有与企业级数据库相媲美的特性，如完善的SQL标准支持、多版本并发控制、时间点恢复、表空间机制、异/同步复制、嵌套事务、在线热备份、一个复杂的查询优化器、预写日志容错技术。
- PostgreSQL 官方网站（https://www.postgresql.org/）

关于PostgreSQL

技术特性

- **开放特性**
  - PostgreSQL内置了丰富的数据类型，如任意精度的数值、无限制长度的文本、几何图元、IP地址、数组等。同时还允许用户定义基于正规SQL类型的新类型
- **可编程性**
  - PostgreSQL同样拥有大量的编程接口供用户开发使用，如ODBC、JDBC（Java）、Libpq（C/C++）等。
- **可定制性**
  - PostgreSQL拥有广泛的编程语言支持来实现函数功能，包括内置的PL/PGSQL过程语言，PL/Perl、PL/PHP、PL/Python、PL/Ruby、PL/Tcl等脚本语言，以及Java、C/C++等高级编程语言。

关于PostgreSQL

技术特性

- **索引手段**
  - 用户可以自定义索引方法或者使用内置B-Tree索引，Hash表索引，GiST索引，GIN索引。
- **多种身份认证方式**
  - PostgreSQL中可以使用数据库用户/角色、操作系统、PAM、Kerberos等方式，根据配置文件（pg_hba.conf）中的设置执行对应的身份认证。
- **支持多平台**
  - PostgreSQL 可以在主要的操作系统上运行，包括：Linux 操作系统(32/64) Windows(32/64) 和 UNIX（包括AIX，BSD，HP-UX，SGI IRIX，Mac 的 OS X，Solaris， TRU64)

关于PostgreSQL

PostgreSQL 无处不在

国际方面用户包括:

- 微软 Microsoft
- CitusDATA
- 亚马逊 Amazon
- 苹果公司 Apple
- 澳大利亚电信公司 Telstra
- 巴斯夫公司 BASF
- 红帽 Hed Hat
- 思科公司 Cisco
- 美国航空航天局 NASA,
- …甚至包括国际太空站 The International Space Station

PostgreSQL 全球财富1000的用户

Accenture 埃森哲
ADP
Aetna 安泰保险
AT&T
AutoZone
BAE Systems
Banco do Brasil 巴西银行
Boeing 波音公司
Bouygues Telecom
Broadcom 博通公司
Cisco Systems 思科公司
Citigroup 花旗集团
Cognizant Technology Solutions
Computer Associates
Computer Science Corporation
Deere & Company
Dell 戴尔公司
Deutsche Boerse AG 德国证券交易所
eBay 易趣网
EMC Corporation 易安信公司
Emerson Electric
Ericsson
Fujitsu 富士通公司
General Electric (GE) 通用公司
Google 谷歌公司
Grupo BBVA
HP 惠普公司
IBM 公司

ICICI
Infosys 公司
JPMorgan Chase
KDDI
KT 韩国电信
Kubota
Kyocera
Lockheed Martin

MasterCard International
McKesson
Mizuho Information & Research Institute
Mosaic ATM
Motorola 摩托罗拉公司
NEC 日电公司
NTT 日本电信
Nokia 诺基亚

Northrop Grumman
Nucor
ONGEI
Panasonic 飞利浦
QUALCOMM
Raytheon
RSA
SAP 公司
Schneider Electric
Seagate 希捷
Siemens 西门子
SK Telecom
Softbank 日本软件银行
Sony 索尼公司
Swisscom 瑞士电信
Symantec 赛门铁克
Syngenta Crop Protection
Tata Consultancy Services
Telstra 澳洲电信
The GAP 时尚服饰
Tokio Marine & Nichido Fire Insurance
Toyota
Union Pacific Railroad
VMWare 公司
Walt Disney
Wipro
Xerox
Yahoo 雅虎公司

关于PostgreSQL

PostgreSQL 无处不在

国内也越来越多知名企业在应用PostgreSQL:

- 工商银行
- 中国邮政储蓄
- 平安集团
- 苏宁
- 京东
- 去哪儿网
- 高德地图
- …等等

基于PostgreSQL相关的数据库产品也不断地在出现:

- 腾讯Tbase
- 阿里PolarDB
- 瀚高HighgoDB
- 亚信AntDB
- 华为OpenGauss
- 金仓KingBase
- 美创MCDB
- …等等

关于PostgreSQL

培训认证，政策导向

- 自从2019年“中国PostgreSQL培训认证”获得工信部下属中国软件行业协会及中国电子工业标准化技术协会权威认证。

- 多家培训机构也纷纷出现成为授权合作伙伴，例如：
  - 恩墨学院
  - 晟数学院
  - CUUG优技
  - 东方瑞通
  - 博森瑞
  - 柯普瑞等

- 在当前的国际竞争格局下，不能掌握核心技术就会受制于人，这对于所有民族基础软件企业、技术从业者都提出了更高要求。

- 在新时期信息安全新形势下，国内大量企业投身于利用以PG为代表的开源技术创新ICT产业的事业中。

- 并在国内造就大量的PG技术人才需求

北京大学 PEKING UNIVERSITY

瀚高 HIGH GO

透明数据加密和密钥管理

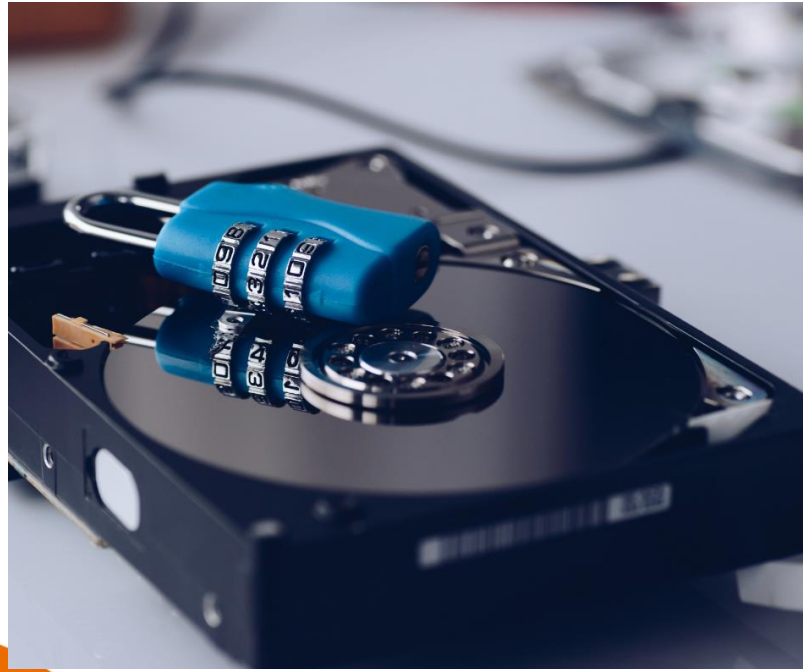
透明数据库加密和密钥管理

只有更好，没有最好

- 既然当前的PostgreSQL已经如此完美了，我们还能做些什么呢？
- 在当前PostgreSQL国际社区里，只有更好，没有最好。
- 当前的PostgreSQL就欠缺了透明数据库加密和密钥管理的功能
- 某些公司或是行业合规性的要求（compliance），尤其是牵扯到国家安全的企业，没有这些安全的功能是不能部署PG的服务。
- 透明数据库加密和密钥管理的功能也是我们此次课程的重点。

透明数据库加密

Transparent Data Encryption（TDE）



- 和网络数据传输的加密有所不同
- 指的是PostgreSQL在存储数据到磁盘上之前做的加密的动作
- 以及在读取磁盘上的加密数据之后做的解密的动作
- 透明，意思是PostgreSQL应当在后台自己把加密和解密的动作都自动完成，用户本身不需要知道具体的细节
- 用来保护存在磁盘上的用户数据，即便黑客可以访问磁盘，但依然无法访问加密数据。

密钥管理系统

Key Management System（KMS）

- 有了加密和解密，就必然涉及到密钥的管理
- 对于密钥的管理及保护又是另一门学问
- 得到了密钥，那基本上就可以解密加密过的数据
- 因此，真正保证数据库的安全，并不仅限于加密和解密的动作。
- 对密钥的保护以及管理是数据库安全比不可少的一个重要环节
- 例如，密钥长度？如何替换？谁可以更换？如何安全的储存密钥等等的一系列问题。

透明数据库加密和密钥管理

参考资料

- 数据库加密和密钥管理在后继的课程会有更详细讨论
- 当前仅需要了解如下基本概念：
  - PostgreSQL(PG)
  - 透明数据库加密（TDE)
  - 密钥管理系统（KMS)

- PostgreSQL 官方网站：（https://www.postgresql.org/）
- PostgreSQL 官方 TDE 和 KMS wiki page：（https://wiki.postgresql.org/wiki/Transparent_Data_Encryption）
- 基本PostgreSQL 指南（https://www.postgresqltutorial.com/）

PostgreSQL内部架构（http://www.interdb.jp/pg/）

透明数据库加密和密钥管理

国内参考资料

- 中国PG分会官网：（https://www.postgresqlchina.com）
- PostgreSQL 官方技术问答：（www.pgfans.cn）
- 资源下载（www.postgreshub.cn）
- PG分会微信公众号：开源软件联盟PostgreSQL分会

HIGH GO

融知与行 瀚且高远

THANKS