

PostgreSQL生态价值与 瀚高开源社区运营实践

丁治明

中科院软件所研究员
瀚高数据库首席科学家

1. PostgreSQL生态价值与瀚高开源战略

第十六届开源中国开源世界高峰论坛

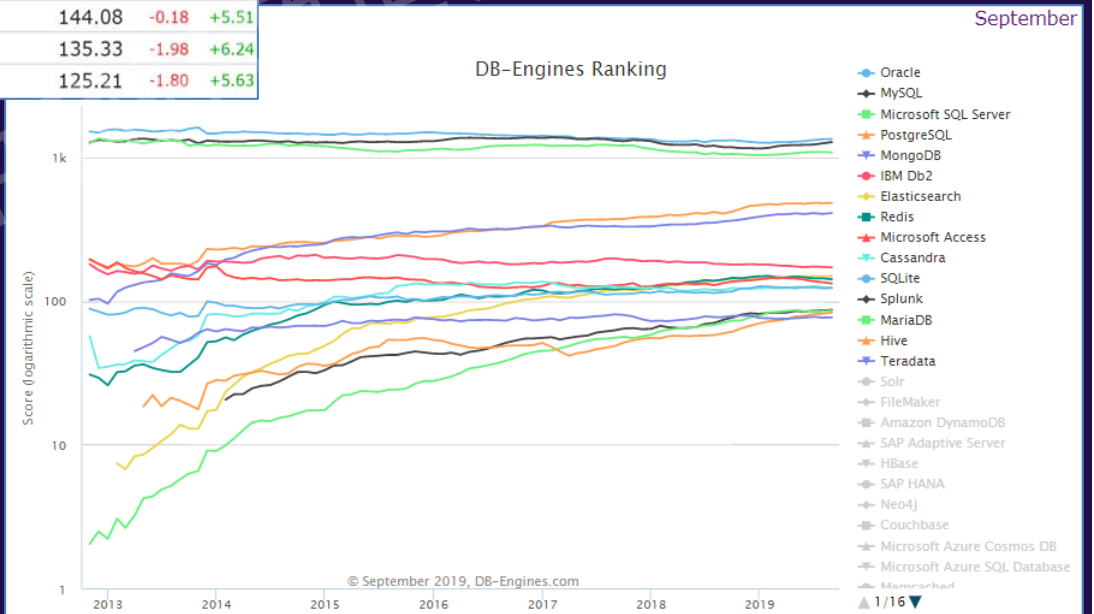
351 systems in ranking, August 2019

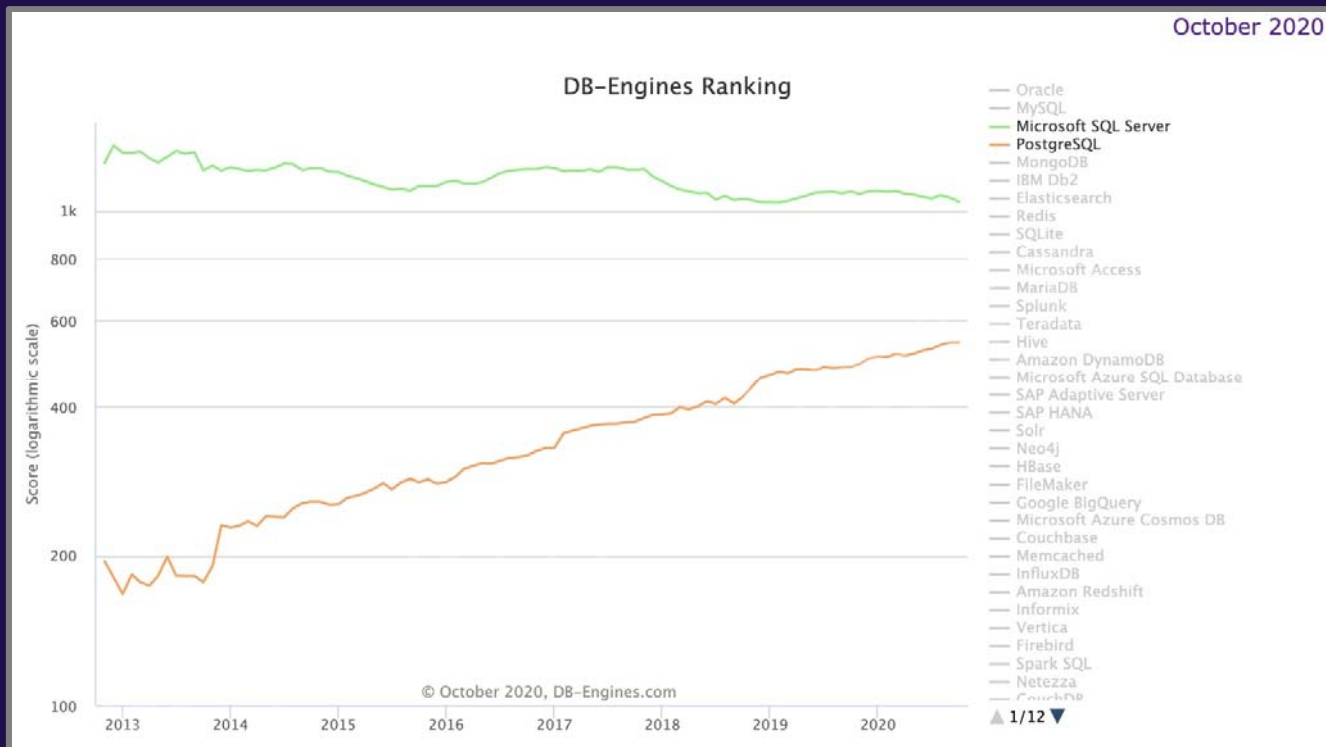
Rank			DBMS	Database Model	Score		
Aug 2019	Jul 2019	Aug 2018			Aug 2019	Jul 2019	Aug 2018
1.	1.	1.	Oracle +	Relational, Multi-model	1339.48	+18.22	+27.45
2.	2.	2.	MySQL +	Relational, Multi-model	1253.68	+24.16	+46.87
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model	1093.18	+2.35	+20.53
4.	4.	4.	PostgreSQL +	Relational, Multi-model	481.33	-1.94	+63.83
5.	5.	5.	MongoDB +	Document	404.57	-5.36	+53.59
6.	6.	6.	IBM Db2 +	Relational, Multi-model	172.95	-1.19	-8.89
7.	7.	↑8.	Elasticsearch +	Search engine, Multi-model	149.08	+0.27	+10.97
8.	8.	↓7.	Redis +	Key-value, Multi-model	144.08	-0.18	+5.51
9.	9.	9.	Microsoft Access	Relational	135.33	-1.98	+6.24
10.	10.	10.	Cassandra +	Wide column	125.21	-1.80	+5.63

关于PostgreSQL

PostgreSQL (简称PG) 是最先进的开源关系型数据库, 世界排名第四, 近年呈迅速上升趋势。PG自80年代伯克利大学Michael Stonebraker(图灵奖获得者), 创造的POSTGRES软件包发展而来, 多年来保持稳定迭代。

不为任何商业实体掌控的PostgreSQL拥有开源领域最具全球性核心贡献者社区, 社区稳定运行超过25年, 构建出无以伦比的社区生态。





对比SQL Server和 PostgreSQL 的增长曲线可以清晰的看到：

PostgreSQL 的增长从未放缓，如保持现在的增长趋势，PostgreSQL 会在 4~5 年内完成对 SQL Server 超越，进军全球数据库领军者阵营

PostgreSQL的生态稀缺性

开放的BSD许可证+ 社区性开源。让PostgreSQL具有永久性安全保障。PostgreSQL具有安全合规，不受美国技术限制的特点，是公认的产权公有的全球性技术。

对于中国来说，PG是为数不多的**可以放心发展的开源技术（供应链可信）、国际性技术、关键基础技术**，是中国加速关键产业发展，强化基础布局的重要路径。正因为此，**PostgreSQL技术的生态稀缺性**逐渐彰显。

瀚高开源社区战略

- 瀚高开源社区战略思想是建立在对开源国产化深层次认知的基础之上，是深刻理解与探索国产基础软件发展正确路线的实践产物。

发展国产基础软件，必须思考三个问题： 瀚高开源社区战略分三个层次，

- 是否是国产数据库发展的根本道路 (why) ?
 - ① 国产数据库开源路线的唯一性；
 - ② 国产开源基础软件发展最优路径选择；
- 是否找到这条道路 (方向) 的最优路线 (what) ?
 - ③ 全球化+国产化开源基础软件社区发展实现路径探索
- 谁能做？怎么做 (how) ?

为什么说“开源国产化是中国数据库发展的必经道路”？

- 基础软件不是“写”出来的，而是“用”出来的，基础软件需海量场景、海量用户的支撑。
- 闭源关系型数据库发展的时间窗口已经过去30年了，国际巨头已建立基础软件生态壁垒，先发优势巨大。
- 传统商业数据库推广模式效率低，国产厂商核心工作在“销售”，工程师的创造性被压制。
- 软件不同于一般实物性商品。闭源模式下重复造车的仿制思维，创新格局不足，弯道超车中无法达到超越巨头的加速度。

国产开源数据库的发展最优路径？

开源路线

- 原生自主开源。 ——难度大，风险大，要求高。
- 分裂式开源 ——难度大，且有悖开源价值观，并有法律风险。
- 跟随式开源/闭源 ——发展慢，自主化低，且违反开源价值观，并有法律风险。

开源样本类型

- 商业开源 ——有核心风险，GPL类自由软件风险更大。
- 社区/学术开源 ——优先选择

厂商类型

- 云厂商 ——基因制约其必须追随各种主流数据库，难以发展核心产品
- 数据库开源创企 ——开源路径选择不科学，或解决小问题，或跟随式开源。
- 传统国产闭源厂商 ——仅开源一些小组件，工具等，缺乏深度思考。

国产开源基础软件发展最优路径

- **站在巨人肩上（消化吸收）：**
选择具有生态稀缺性的，全球性、社区性关键开源技术——PostgreSQL。
蛰伏与沉静，对巨人的理解和消化需要漫长的沉淀。
- **巨人为我所用（瞄准中国市场，建立国际化协作社区）**
以中国市场需求为蓝本，以国际化合作创建新型开源协作社区
闭门造车不可为——关系型数据库发展必须有国际力量的参与和指导。
- **降服巨人（发展核心贡献者，左右开源技术的发展）**
发展中国的核心贡献者团队，对国际主流技术社区进行内部洗牌。
基础软件的成功，其标志是杰出贡献者的大量涌现。
- **巨人再成长（发展用户生态，发展中国标准）：**
引导社区开源挑战传统数据库价值市场。以开源模式颠覆OLTP产业格局。

技术创新，实现超越

软件产品有应用才有锤炼，有需求才能迭代，
有市场才会有研发的持续性投入



应用创新，国产化生态完善

数据库是基础软件，底层依赖于CPU、操作系统，
上层与中间件、业务系统紧密结合。生态的完善需要大量人才支撑



满足基本需求，存量系统替代

仍然存在大量IOE系统支撑的业务系统，需要能够迁移到国产化系统平台



完善生态 — 开源产品社区生态





实现超越 — 开源底层创新

国内数据库软件行业如果仅依靠国家政策、低成本运营、本土化营销等，很难挑战国外品牌、建立真正的竞争力；“信创”行业为国内厂商提供了充分的市场机会，只有从底层创新做起，结合国内需求、结合技术发展趋势、融入整个产业链，才能实现真正意义上的超越。

3. 中科院软件所、瀚高数据库开源运营成果

第十六届开源中国开源世界高峰论坛

瀚高开源运营成果

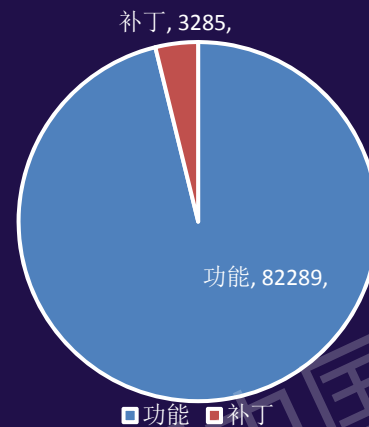
1. 与国际社区核心组保持着密切沟通，承担国际交流桥梁作用。
2. 与国际知名PG企业（EDB、CrunchyData、NTT、SRA OSS、Percona等）保持密切技术交流与国际合作，如与EDB、NTT等共同合作研发PG-POOL II系统等。
3. 深度参与国际社区协作，2020年位列社区顶级贡献者团队。
4. 组织PostgresConf亚洲会议（PostgresConf全球顶级会议），在中国及亚洲组织相关的国际PG会议，推动PG的发展和国内生态的建设。
5. 为社区提交PG代码（数万行），并协助测试，从而协助提升PG社区所提及的代码质量。
6. 与PG全球官方网站PostgreSQL.ORG维护组织有密切沟通，PG历年发行版瀚高除提供各类代码贡献之外，还编写了PG V8-V13的中文发行说明和产品文档。

国际社区贡献

亚洲区域最大的贡献力量

代码提交：

- 功能提交 (82289行)
- Patch提交 (3285行)
- Bug反馈 (60+)
- 目标：开源社区核心贡献者

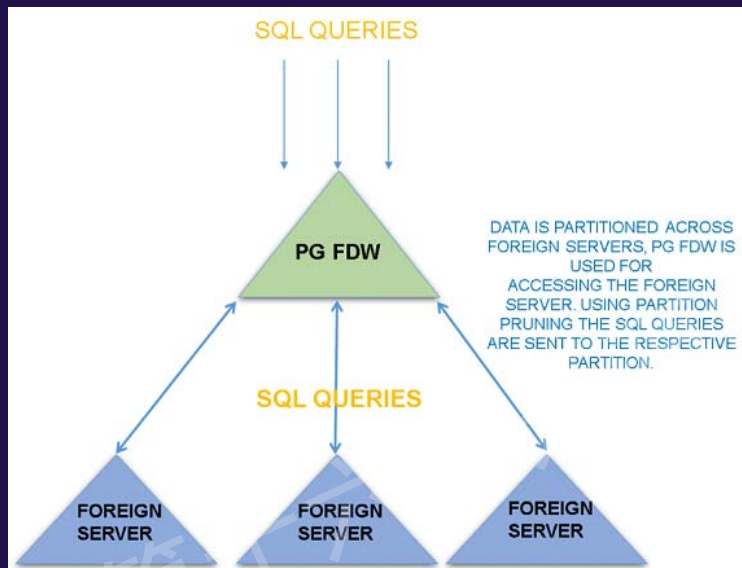


社区活动：

- 中国开源软件推进联盟PostgreSQL分会发起者与秘书长单位；
- 国内历年大象会主要参与单位，会议组织、话题演讲；
- 组织PGOPEN CHN 2017-19，发起成立国际顾问委员会；
- 2017年开始，参加PostgreSQL历届国际开发者会议；
- 2019年PGConf US主要赞助单位之一，国内唯一赞助商；PGConf CN（北京）；
- 2020年PGConf CN与PGConf ASIA合办，主要组织者



开源研发工作（一）：Build-in Sharding



- Shard Management : 分片管理、DDL下推 ;
- Global transaction manager : 两阶段提交 ;
- Parallel foreign scan : 分布式跨节点查询性能 ;
- Global snapshot : 保障分布式事务跨节点原子性可见性
 - Global CSN based snapshot : 基于时间戳的全局CSN实现 ;
- Bulk DML Operations : 跨节点数据增删改效率
- Cluster Management : 备份、恢复、监控、高可用等工具

开源研发工作（二）：数据加密

访问控制层，防止非授权用户访问

- 以角色为基础（三权）
- 自主访问控制（DAC）
- 强制访问控制、多级安全

数据加密，防止不可信的数据库服务者

- 确保数据库索引等功能正常使用
- 不解密情况下对数据正常操作
- 大数据量已加密数据的密钥定期更改



数据文件加密，防止数据文件窃取

- 全数据库文件加密
 - 表级数据文件加密
- 对数据块进行加密，IO过程中完成加解密操作，采用XTS加密模式，支持国密算法sm4

开源研发工作（三）：pgpool II

- 新版本计划：按需生成子进程，提高性能和资源利用率
- 为管理pgpool II集群构建第二代GUI工具
- 在流复制（streaming replication）中将特定后端升级为主数据库
 - -- 添加新参数“backend_promote_priority”以指定提升优先级。
 - -- 发生故障转移时，具有最高优先级的节点将去到failover_command
- 当allow_sql_comments = on启用时，允许Pgpool-II忽略/* NO LOAD BALANCE */注释
- 不计算监视工具中的空查询

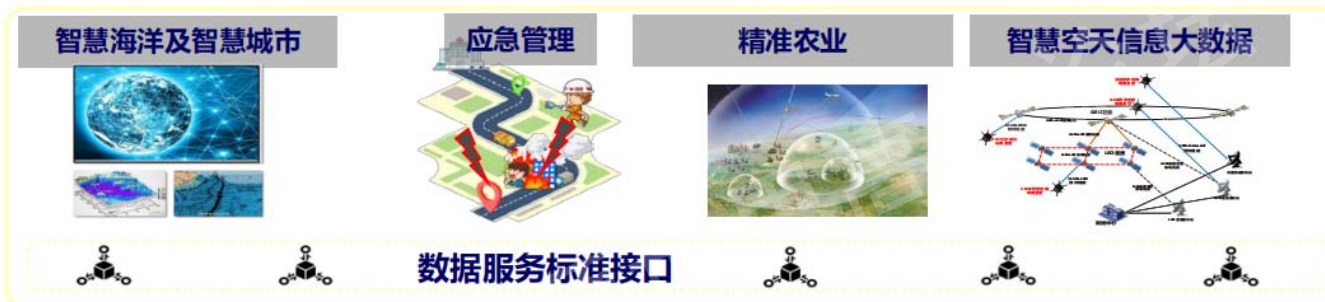
中科院软件所与瀚高数据库的研发合作

双方共同开展基础软件“卡脖子”技术研发。共建联合实验室。开展“国产化高并发分布式数据库系统关键技术研究与应用”等重大关键技术攻关，开发具有自主知识产权的国产分布式数据库系统。

面向国产化平台，合作研究高并发情况下分布式事务管理技术，对分布式事务处理协议优化；深入攻关：分布式数据操作的并行拆分及并行查询优化技术、分布式系统中的Hybrid Logical Clock技术、分布式数据存储动态调整技术；及研制分布式系统数据多副本管理、备份、监控等配套工具。

中科院软件所与瀚高数据库合作，共同争取“山东数据科学产业研究院”获批。发展时空数据管理前沿技术及关键应用。

领域支撑共性
服务及关键应用



时空大数据基础设施与数据中心



时空大数据
信创基础软件



第十

CCOW
OPEN SOURCE CHINA
OPEN SOURCE WORLD

The 16th
Open Source China
Open Source World Summit

第十六届开源中国开源世界高峰论坛
Embrace Open Source Software, Drive Global Innovation
拥抱开源 缔造创新模式

3. 中国开源软件推进联盟PG分会工作实践

第十六届开源中国开源世界高峰论坛

行业协会-中国开源软件推进联盟PG分会

性质：在2000年初即诞生的PG中文社区基础上组织成立的PG中国行业协会组织。基于PG的影响以及国内数据库技术的发展现状，为推动国内PG产业生态建设和产学研用发展，加强国内开源技术建设和应用，促进国产基础软件发展。

成立：2017年，在工信部中国电子信息产业发展研究院指导下，依托中国开源软件推进联盟，特成立PostgreSQL分会（简称中国PostgreSQL分会或PG分会）
（官网：www.postgresqlchina.com）



成立仪式：

Deb Goodkin（FreeBSD基金会执行董事）

Jim Zemlin（Linux基金会执行董事）

陆首群（前国务院信息化联席会议办公室常务副主任、中国开源软件推进联盟名誉主席）

苗健（瀚高软件董事长）

刘澎（中国开源软件推进联盟副主席兼秘书长）

汪洋（业内知名专家、平安科技数据库负责人）

中国PG分会工作内容

PGConf.CN/Asia亚洲大会
PG全球顶级会议

1

4

PostgreSQL高校讲座及学术会议
与CCF合作，面向国内高校的
PostgreSQL推广及学术交流会议

Postgreshub中国开发者社区
中国PG开发者社区和PG中国开
源社区版产品

2

活动
与
平台

5

代表中国参加PostgreSQL国际会议
促进中国与全球PG组织的沟通、交流
与合作

PGFans用户社区
Postgre.Live国际沙龙
全球唯一PG中文技术问答社区

3

6

PostgreSQL中国培训认证体系
中电标协、中国PG分会联合认证

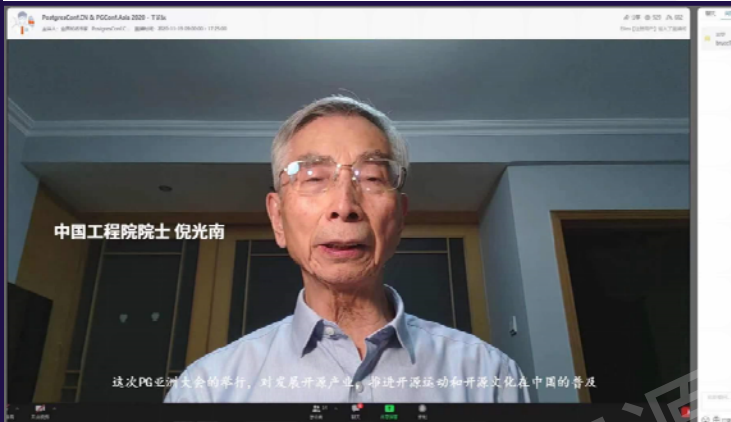
PGConf.CN中国大会/.Asia亚洲大会

PostgresConf是全球最大PG会议序列。PostgresConf.CN大会由PG分会与PostgreSQL全球基金会、PGConf国际会议组织联合举办，是中国规格及影响力最高的PostgreSQL技术及产业大会。

PGConf.Asia大会是亚洲最顶级的PG技术大会，**参会国际演讲嘉宾超过50位**。2020年度首次落户中国，并与PostgresConf.CN2020大会联合举办。**是大陆区域国际嘉宾最多的IT会议之一**。



PCC POSTGRESCONF CN 2020 | PGConf.Asia



PostgresConf.CN & PGConf.Asia 2020

生而全球·PG在中国乃至亚洲的生态价值

PostgresConf.CN & PGConf.Asia 2020



亚洲规模最大的 11.17/18/19/20
开源关系型数据库大会 全球独家直播

PostgresConf.CN联合PGConf.Asia, 携手100多位Top企业家, 科学家, 资深大咖为你打造一场持续4天的知识狂欢盛宴, 104场演讲, 5场专场培训, 14个直播频道。



PostgresConf.CN & PGConf.Asia 2020



亚洲规模最大的 11.17/18/19/20
开源关系型数据库大会 全球独家直播

PostgresConf.CN联合PGConf.Asia, 携手100多位Top企业家, 科学家, 资深大咖为你打造一场持续4天的知识狂欢盛宴, 104场演讲, 5场专场培训, 14个直播频道。



PostgresConf.CN & PGConf.Asia 2020



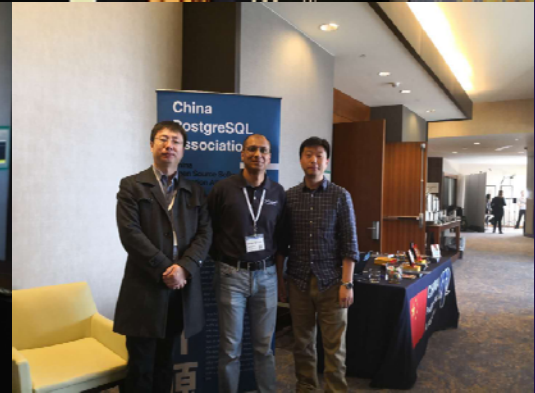
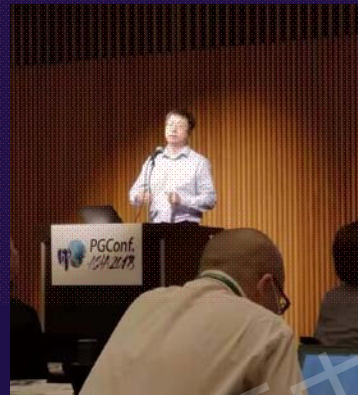
亚洲规模最大的 11.17/18/19/20
开源关系型数据库大会 全球独家直播

PostgresConf.CN联合PGConf.Asia, 携手100多位Top企业家, 科学家, 资深大咖为你打造一场持续4天的知识狂欢盛宴, 104场演讲, 5场专场培训, 14个直播频道。



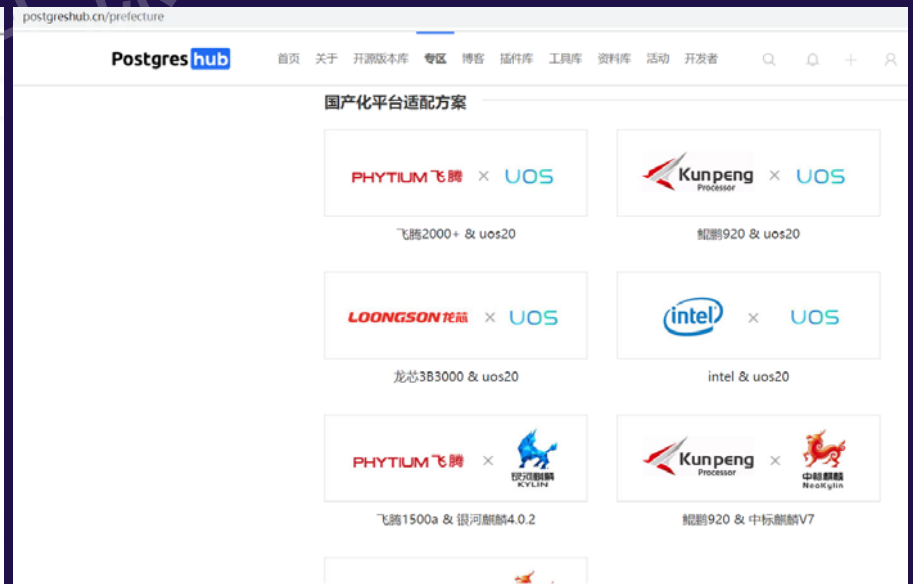
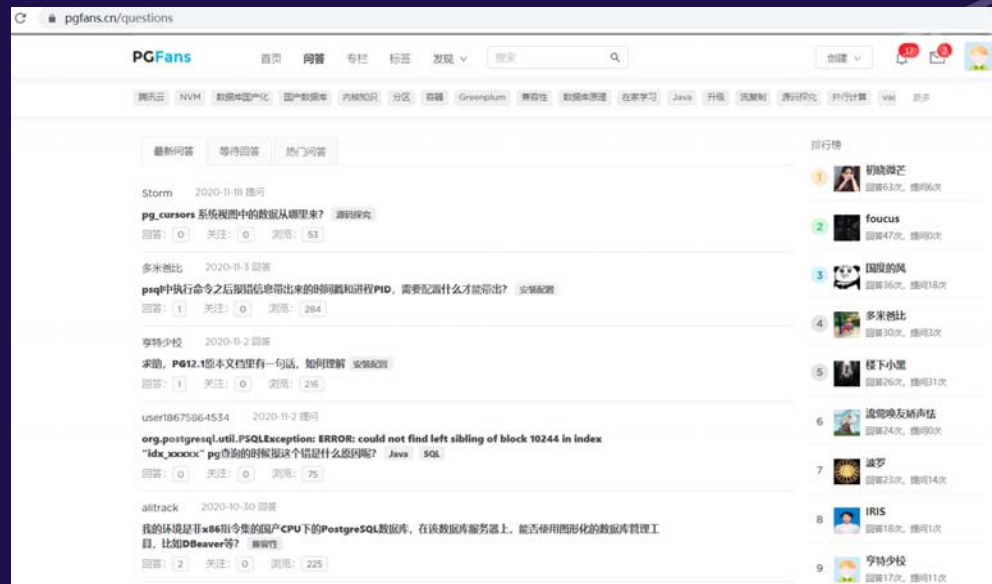
代表中国参加国际PG会议

PGconf US ,
PGCON 加拿大 ,
PGconf EU
PGVision US ,
PGCONF Asia等



PG中国开发者社区、用户社区线上平台

- www.pgfans.cn
- 全球唯一的PostgreSQL中文技术问答社区
- www.postgreshub.cn
- PostgreSQL中国开发者社区的官方平台



国内唯一的PG技术培训和能力认证体系。

知名培训机构和专业讲师
承接具体的培训服务

培训

中国PG分会
培训认证体系

受众

PG从业者
参加技术培训和能力考核
学有所获，证有所有

中国电子工业标准化协会
联合提供权威技术能力认证
使学习成果体现为能力证书

认证

内容

业内资深技术专家
根据多年从业经验，
构建培训大纲和认证考核大纲



PGConf.CN 2019大会正式发布

考核现场



The 16th
Open Source China
Open Source World Summit

第十六届开源中国开源世界高峰论坛
Embrace Open Source Software, Drive Global Innovation
拥抱开源 缔造创新模式

中国PG培训认证体系 — 除Oracle外最具有教育产业化可行性的DB培训体系



中国PostgreSQL技术能力认证考试—博森瑞01考场

中国PostgreSQL技术能力认证考试—北京云和恩墨考场



中国PostgreSQL技术能力认证考试—晟数学院01考场

中国PostgreSQL技术能力认证考试—标准化考场



PostgreSQL Certified Exam
中国PostgreSQL认证考试

中国开源软件联盟PG分会
中国电子工业标准化协会
中国软件行业协会



中国PG培训认证产品生态：

瀚高HGDB，腾讯云Tbase，中兴GoldenDB，亚信AntDB，文武EcoX



PostgreSQL Certified Exam
中国PostgreSQL认证考试

中国开源软件联盟PG分会 | 中国电子工业标准化协会 | 中国软件行业协会

PostgreSQL培训认证国产化产品生态



中国PG分会产教融合工作

DBMS数据库系统实现课程

数据库教授团队



信创学院、信创产教联盟

山东省信创产教联盟

开源数据库创新实验室
PG高校实训基地

数据库/信创实训基地



PostgreSQL
校企合作



课改支持、精品课共建

北大、清华、武大、深大、山师大等
十余所高校

科研合作，学术研究
国家特色示范软件学院等

清华、中科院、郑州大学、山大战略合作



共建专业、人才培养、就业服务

腾讯、阿里、华为、平安、苏宁、工
行、邮储、百度、神州数码

PostgreSQL高校课改计划

“高校数据库系统实现实践教学课程改革”计划：

北京大学、清华大学、武汉大学、山东大学、中国海洋大学、深圳大学、沈阳航空航天大学、山东师范大学、青岛大学、北京科技大学.....



第十六届开源中国开源世界高峰论坛

CCOW
OPEN SOURCE CHINA
OPEN SOURCE WORLD

The 16th
Open Source China
Open Source World Summit

第十六届开源中国开源世界高峰论坛
Embrace Open Source Software, Drive Global Innovation
拥抱开源 缔造创新模式

谢谢 Thanks...

第十六届开源中国开源世界高峰论坛