

# 为什么 PG 这么火

阿里云

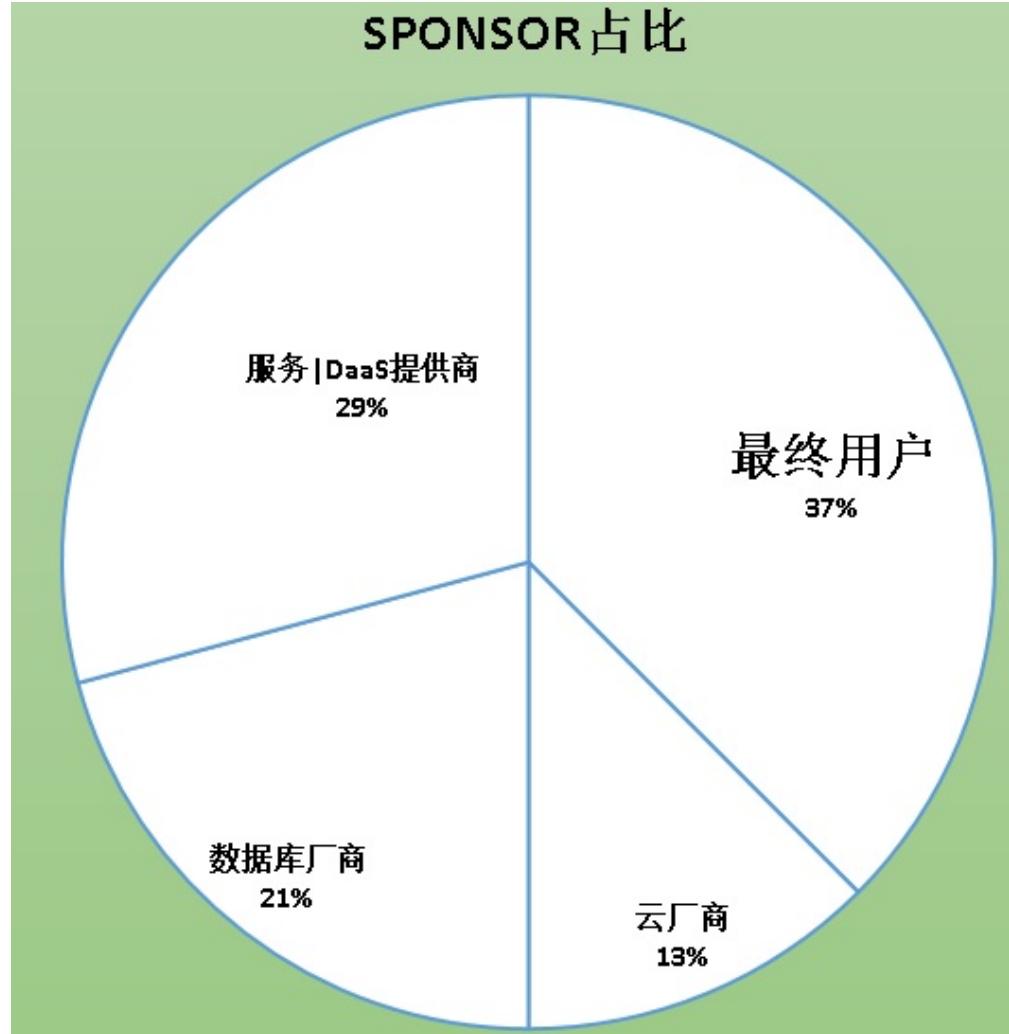
digoal

# 目录

- PG 是什么
- PG 新版本开发者特性精选
- PG 与 Oracle的全方位对比

PG是什么？

社区组成独特性：  
如果：99%的开源数据库都是商业公司  
那么：PG是那1%



## PG 是 社区 开源 数据库 !



# 为什么要贡献代码，成为sponsor？

- 数据库厂商
  - 推一款新的商业数据库，通常都需要背书，小厂产品，谁为你背书？
  - 1、有技术的厂商，很难挑战已有数据库市场格局。
  - 2、有渠道的厂商，需要抓住窗口期，快速占领市场，避免重复造轮子，需要一款可以无法律风险，二次分发的开源数据库。唯有PG。
  - 贡献核心代码，社区所有的用户都可以为之背书
- 数据库服务 | DaaS服务提供商
  - 开源产品的服务提供商，能力如何体现？
  - 架构能力？优化能力？管理能力？FIX BUG的能力？
  - 贡献核心代码，社区所有的用户都可以为之背书
- 最终用户(不靠数据库赚钱)
  - 希望社区长久，期望可以享受免费的、**可持续发展的**、开源的、不被任何商业公司、不被任何国家控制的企业级的数据库。去O(贵)，去DB2去sybase(产品下市、无法再升级、支持)
  - 1、PG用到人越多，越多人背书，使用越靠谱（事实证明，比商业数据库还靠谱），
  - 2、抛砖引玉，企业投入2个研发持续贡献(一年可能一两百万)，实际上整个PG社区有数千人在贡献，对最终用户来说，简直赚到了。
    - 使用商业数据库，除了LICENSE等成本，依旧需要投管理、研发、外包资源，一年数千万。
    - 公司越大，越有动力去贡献社区（事实证明）。从趋势来看，给PG贡献代码的大客户只会越来越多。
- 云厂商
  - 开源数据库纷纷改协议
  - 数据库市场巨大、自研是必然
  - 自研的问题：需要培养生态，需要市场背书，需要大量研发资源，可能需要重复造轮子？
  - BASE PG，1 免去自己培养生态，2 避免重复造轮子，3 PG 的代码基础非常不错(开源界“O” )
  - 防止其他厂商控制PG失去市场主导能力(AWS,google,IBM,微软 都已进入PG sponsor行列)

# PG是有组织有纪律的开源数据库！



# 开源许可独特性-活雷锋

- 类BSD许可。随意分发、闭源或开源。

PG 是 活 雷 锋 !

## License 目

PostgreSQL is released under the [PostgreSQL License](#), a liberal Open Source license, similar to the BSD or MIT licenses.

PostgreSQL Database Management System  
(formerly known as Postgres, then as Postgres95)

Portions Copyright © 1996-2019, The PostgreSQL Global Development Group

Portions Copyright © 1994, The Regents of the University of California

Permission to use, copy, modify, and distribute this software and its documentation for any purpose, without fee, and without a written agreement is hereby granted, provided that the above copyright notice and this paragraph and the following two paragraphs appear in all copies.

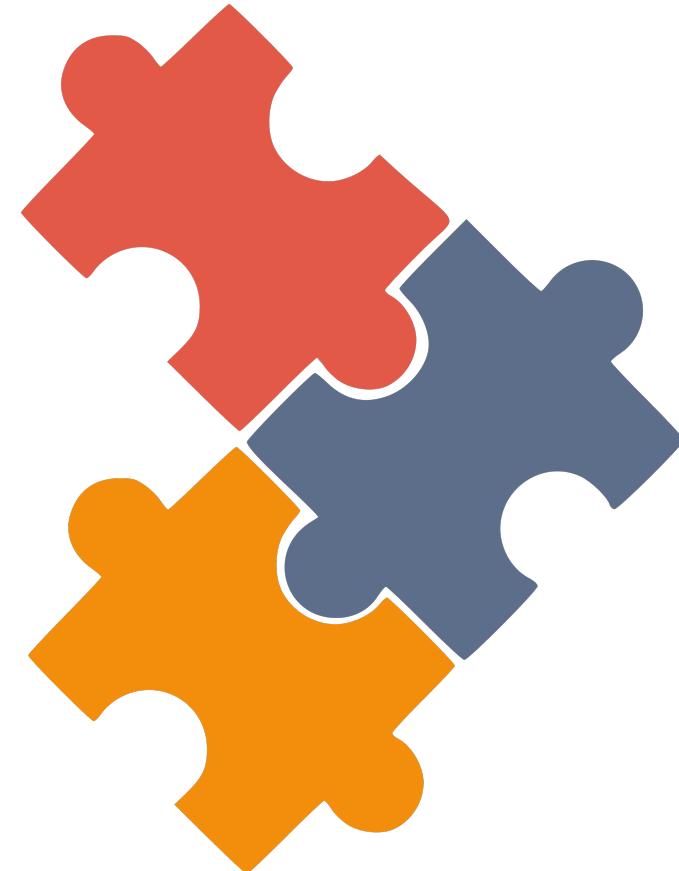
IN NO EVENT SHALL THE UNIVERSITY OF CALIFORNIA BE LIABLE TO ANY PARTY FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES, INCLUDING LOST PROFITS, ARISING OUT OF THE USE OF THIS SOFTWARE AND ITS DOCUMENTATION, EVEN IF THE UNIVERSITY OF CALIFORNIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

THE UNIVERSITY OF CALIFORNIA SPECIFICALLY DISCLAIMS ANY WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE SOFTWARE PROVIDED HEREUNDER IS ON AN "AS IS" BASIS, AND THE UNIVERSITY OF CALIFORNIA HAS NO OBLIGATIONS TO PROVIDE MAINTENANCE, SUPPORT, UPDATES, ENHANCEMENTS, OR MODIFICATIONS.

# 技术架构独特性

PG 是全方位可扩展数据库！

- PG采用了开放接口的设计，如
  - type, operator, index,
  - storage, language, fdw,
  - custom scan, sample, hook 等
- 基于PG开放接口扩展的产品或插件
  - 基于PG的图数据库 agensgraph, edgedb、
  - 流数据库 pipelinedb、
  - GIS PostGIS, pgrouting, pgpointcloud、
  - 时序数据库 timescaledb、
  - 推荐数据库 recdb、
  - 搜索引擎 Yahoo! Everest 、
  - MPP greenplum, redshift , asterdata
  - 机器学习 madlib、
  - 图像识别 imgsmrl、
  - 分词 zhparser, pg\_jieba、
  - 分布式数据库 citus, pg-xc, pg-xl , 高斯 , antdb , polardb
  - Oracle兼容 edb , polardb
  - 文档数据库 torodb
  - 向量搜索，阿里rds pg（图像搜索， 2500万图片约16毫秒）



# 商业趋势

PG 是基石！

- 1、提高安全、合规、正版化意识是全球趋势
- 2、PG是用户去O开源数据库首选
- 3、PG是厂商首选，避免重复造轮子， PG功能覆盖度高，基础好，开源协议友好。
  - 扩展接口多，厂商可以瞄准**行业痛点、差异化输出。**
    - **Citus,tbase,polaradb,pg-xl,antdb,Greenplum,redshift,hashdata,hawq,higo,...**
  - 渠道型厂商，包装PG，快速布局线下输出。
- 4、PG 是云厂商首选
  - 商业开源数据库客户被云收割。商业数据库厂商被迫修改许可协议。
  - **PG是纯社区开源数据库，没有修改协议的动机，是云厂商首选。**

# 技术趋势

PG 是趋势！

- 既要**SQL通用性**，又要**NOSQL扩展性**，还要**多模开发便捷性**。
- 既要**OLTP**又要**OLAP**。
  - 1、PG是多模数据库，开放扩展接口，多模接口丰富。
  - 2、内置并行计算(36种并行计算场景(几乎涵盖所有分析 SQL)，平均提速20倍以上)。
  - 3、PG 12 on ECS tpcc 达到 180万 (tpmc total)
  - 4、支持存储引擎扩展 (12+:
    - heap,
    - zheap,
    - zedstore行列混存储,
    - vops(向量加速存储) )。
  - 5、对芯片友好。

# PostgreSQL 荣誉

PG 是学术界 + 工业界先进代表！

DB-engines  
2017 年度数据库



DB-engines  
2018 年度数据库



2019 OSCON  
终身成就奖



继Linux后第二个获得该奖的开源产品

# PG 定位-企业级、纯社区开源、多模数据库

创新价值

OLTP、OLAP、  
SMP并行计算、  
GPU并行计算、  
列存储、  
JIT、向量计算

PG是企业替换O首选！

混合  
负载

多模

商用价值

0丢失+高可用:  
quorum based sync repl  
安全:  
SSL\SQL防火墙\TDE(13)  
弹性:  
插件化 shard  
模块化扩展  
性能:  
32C TPCC 100万+  
并行计算

Oracle  
兼容

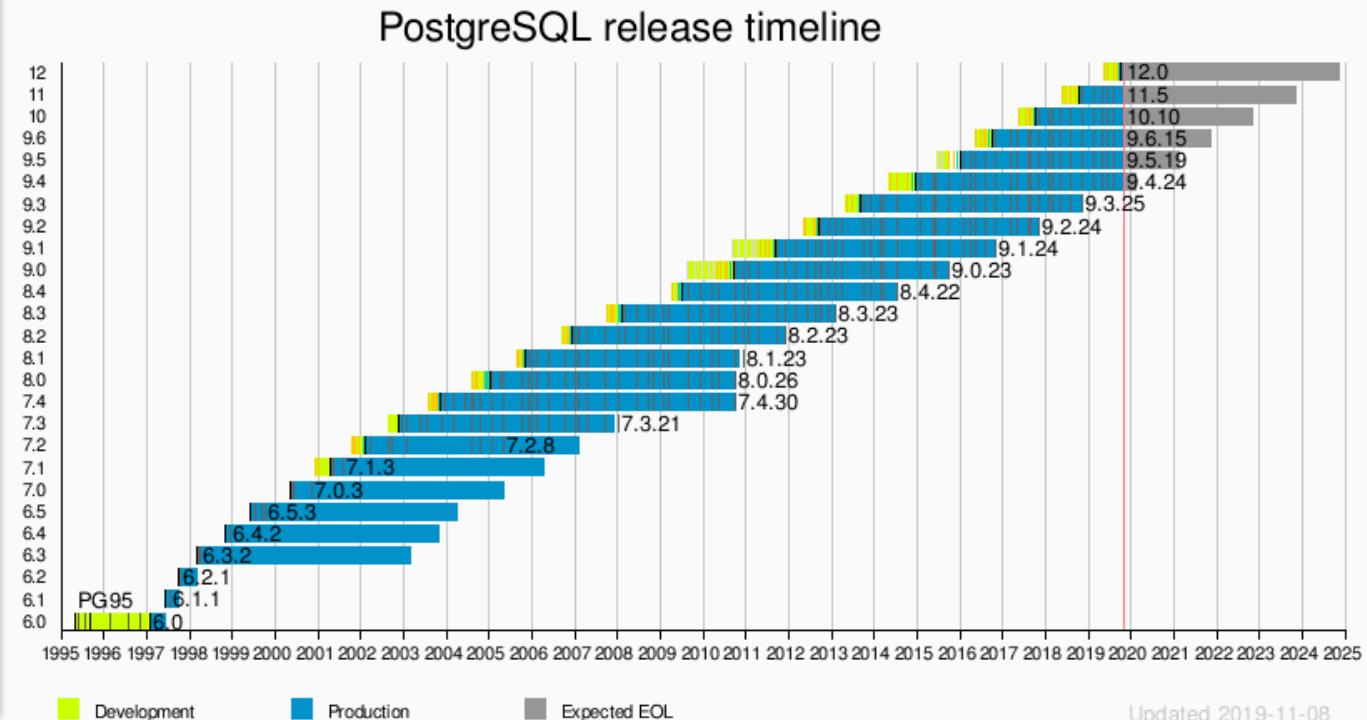
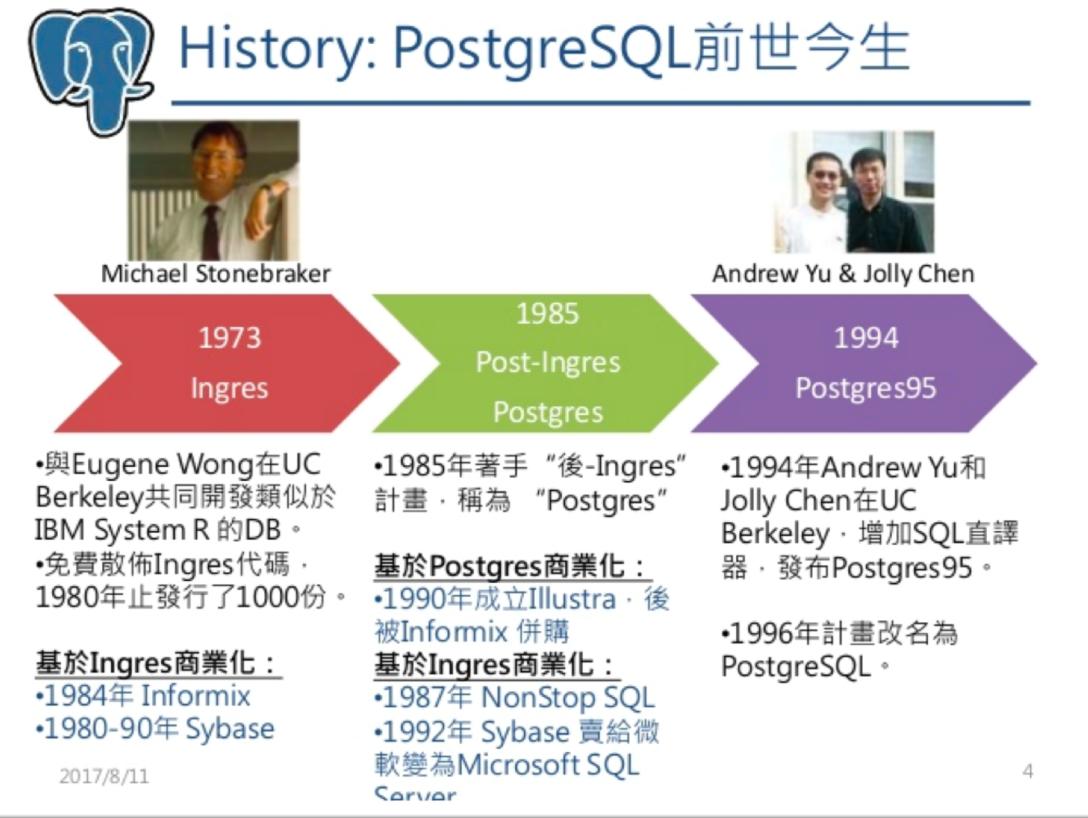
开放接口:  
时空、GIS、图像  
文本、时序、  
向量相似、图谱  
流计算、异构、  
机器学习、  
多维计算、shard

社区版:  
oracle  
阿里云:  
POLARDB for Oracle

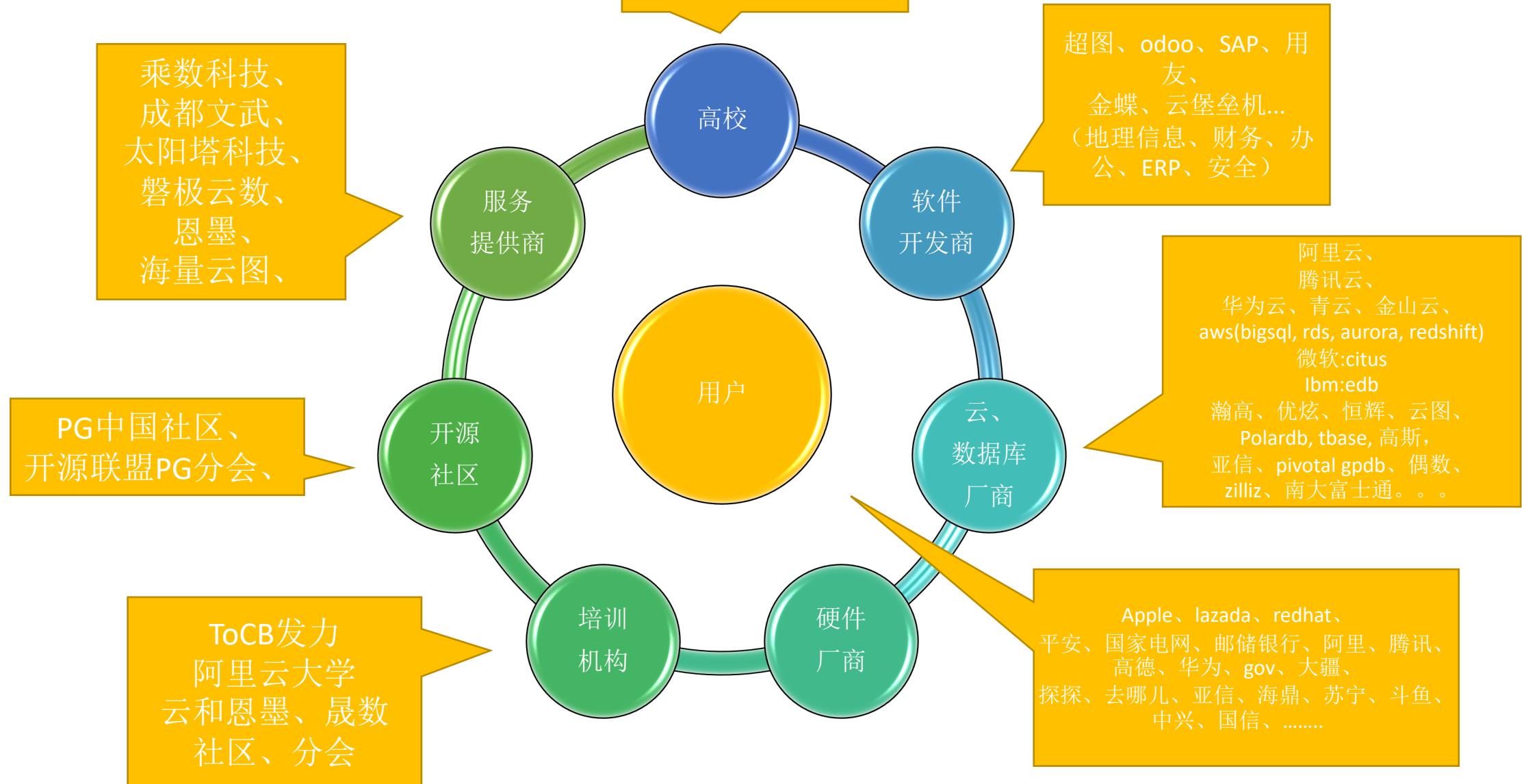
# PG 历史

PG 是常青树！

每年一个大版本，  
小版本平均1.5个月发布，  
每个大版本维护5年。  
历史版本特征矩阵：  
<https://www.postgresql.org/about/featurematrix/>



# PG 是生态！



# PG是有巨大社会价值的开源数据库！

Please note that since PostgreSQL is completely free, and there are no registration requirements, such reporting is completely voluntary, and many places that use it do not advertise that they do. So it's pretty hard to say "here are our top 10 users" when we don't really know who the biggest users are. For all we know the Fortune 50 could be all heavily using it and we'd never know unless someone there spoke up and told us.

PG 衍生数据库、插件	应用
Greenplum、redshift	大数据分析
Agensgraph\edgedb	图计算
asterdata	大数据分析
Citus、xc\xl、antdb	HTAP
timescaledb	IoT时序数据
Edb\POLARDB 兼容Oracle	替换 商业数据库
pipelinedb	流计算
Yahoo everest	搜索引擎
Imgsmrl、palaemon	图像识别、多维搜索
madlib	机器学习

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

PG是什么？

重新定义开源数据库

# 目录

- PG 生态
- PG 新版本开发者特性精选
- PG 与 Oracle的全方位对比

# PG 11

- 并行计算增强
- **btree index include**索引叶子附加属性
- 添加字段（含默认值）更快

# 并行计算增强 - 一库多用



高性能  
几乎线性提升

- 单表10亿测试
- 平均提速20倍+



自动计算并行度

- 不需要改写SQL
- 自动根据SQL代价启动并行计算

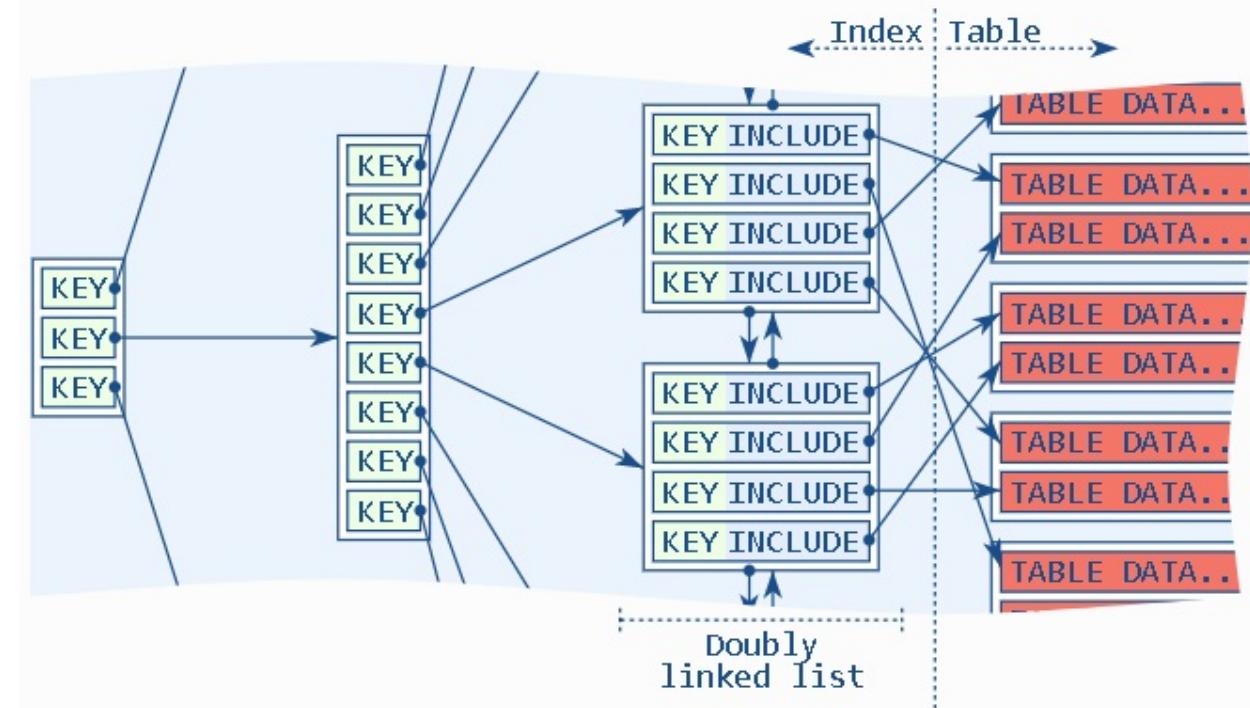
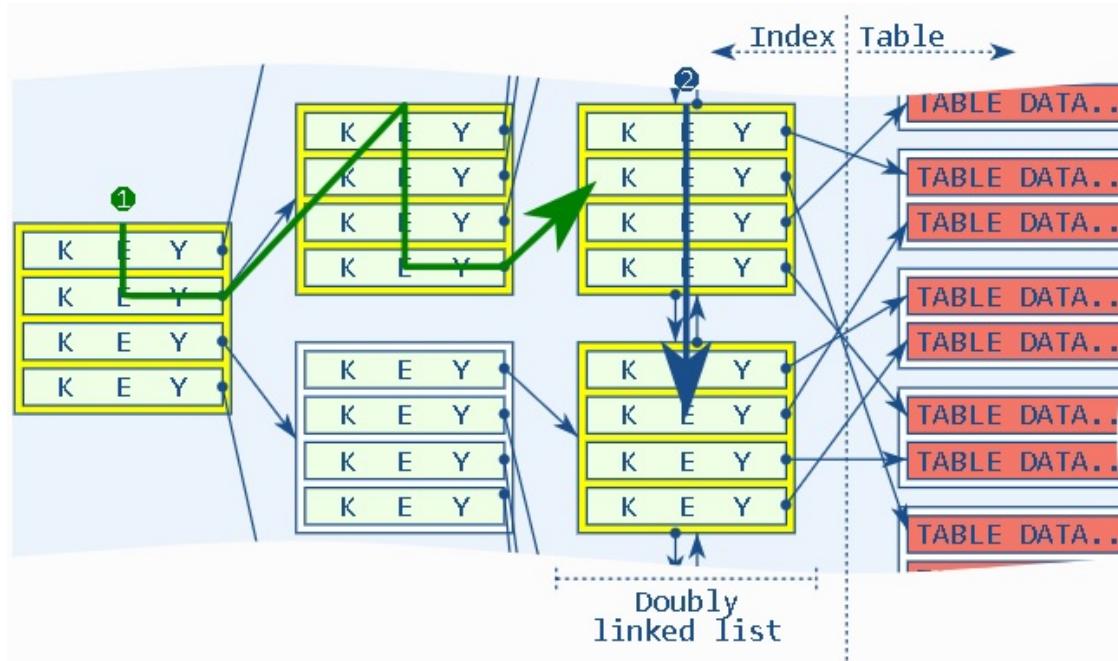


覆盖几乎所有复杂查询

- 表扫描
- 索引扫描
- SQL 过滤
- JOIN、子查询
- 聚合、自定义函数
- 分区表、外部表
- 自定义聚合、函数
- 排序、建表、物化、创建索引

# btree index include

- [https://github.com/digoal/blog/blob/master/201905/20190503\\_03.md](https://github.com/digoal/blog/blob/master/201905/20190503_03.md)
- 类似B+tree数据聚集,
- create index idx\_t1\_1 on t1 (id) **include(c1,c2,c3,info,crt\_time);**
- 痛点:**
- 数据离散存储, 查询多条时IO多, 查询慢, 占用内存多, memcpy效率低



# 添加字段（含默认值）更快

- [https://github.com/digoal/blog/blob/master/201805/20180518\\_01.md](https://github.com/digoal/blog/blob/master/201805/20180518_01.md)

```

postgres=# \x
Expanded display is on.
postgres=# select * from pg_attribute where attrelid='aaa'::regclass and attname='c1';
-[ RECORD 1 ]+-----
attrelid | 99498
attname   | c1
atttypid  | 25
attstatttarget | -1
attlen    | -1
attnum    | 4
attndims  | 0
attcacheoff | -1
atttypmod  | -1
attbyval   | f
attstorage | x
attalign   | i
attnotnull | f
atthasdef  | t
atthasmissing | t
attidentity |
attisdropped | f
attislocal  | t
attinhcount | 0
attcollation | 100
attacl     |
attoptions  |
attfdwoptions |
attmissingval | {digoal}
Time: 0.470 ms

```

```

postgres=# create unlogged table aaa(id int, info text, crt_time timestamp);
CREATE TABLE
Time: 6.259 ms
postgres=# insert into aaa select generate_series(1,1000000),repeat(md5(random()::text),10), now();
INSERT 0 1000000
Time: 2151.531 ms (00:02.152)
postgres=# insert into aaa select * from aaa;
INSERT 0 1000000
Time: 1235.480 ms (00:01.235)
postgres=# insert into aaa select * from aaa;
INSERT 0 2000000
Time: 2688.409 ms (00:02.688)
postgres=# insert into aaa select * from aaa;
INSERT 0 4000000
Time: 4782.437 ms (00:04.782)
postgres=# insert into aaa select * from aaa;
INSERT 0 8000000
Time: 11367.010 ms (00:11.367)

postgres=# \dt+ aaa
          List of relations
 Schema | Name | Type | Owner | Size | Description
-----+-----+-----+-----+-----+
 public | aaa  | table | postgres | 5618 MB |
(1 row)

```

2、新增字段，并添加默认值，由于只需要修改元数据，瞬间完成。

```

postgres=# alter table aaa add column c1 text default 'digoal';
ALTER TABLE
Time: 3.013 ms

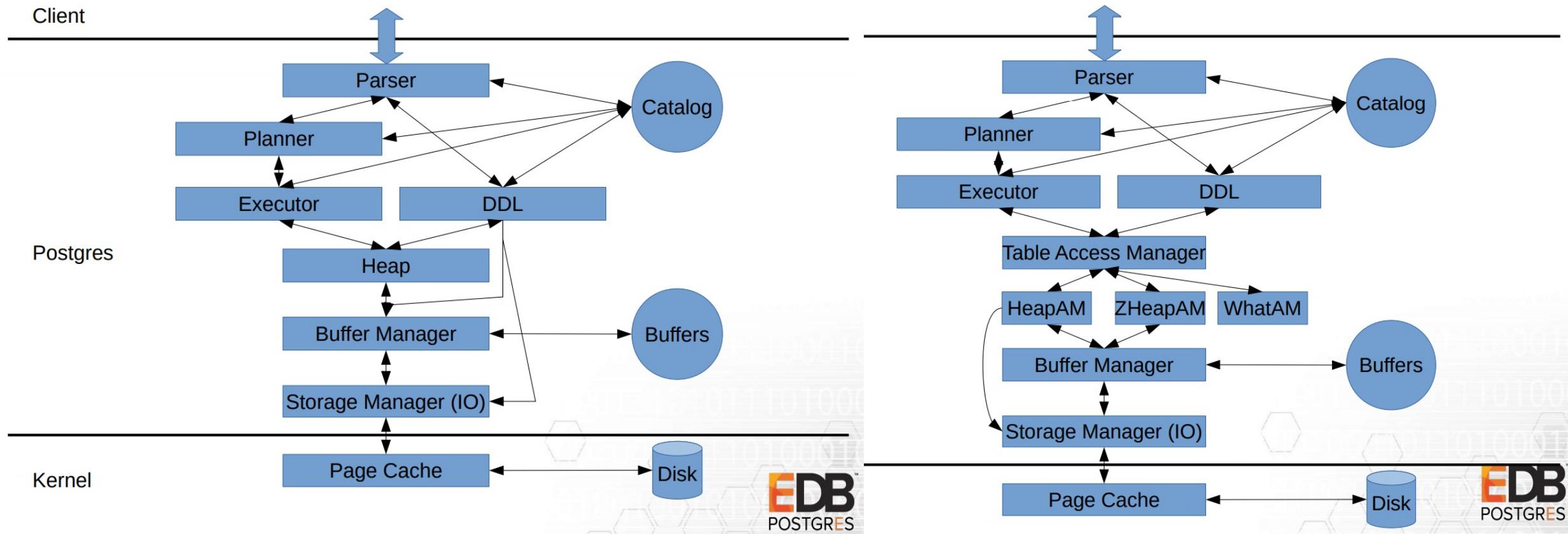
```

# PG 12

- AM 接口
- 分区表大幅性能提升
- GiST index include 索引叶子附加属性
- 日志采样
- 自定义统计信息支持多列 MCV
- Reindex concurrently 不堵塞读写操作
- SQL/JSON path language
- Generated columns
- case-insensitive and accent-insensitive grouping and ordering

# AM接口

- [https://github.com/digoal/blog/blob/master/201903/20190331\\_03.md](https://github.com/digoal/blog/blob/master/201903/20190331_03.md)
- [https://github.com/digoal/blog/blob/master/201905/20190531\\_03.md](https://github.com/digoal/blog/blob/master/201905/20190531_03.md)



- zedstore(列存)
- zheap(支持回滚段)

```
postgres=# SELECT AVG(i199) FROM (select i199 from layout offset 0) x; -- heap
          avg
-----
 500000.500000000000
(1 row)

Time: 4679.026 ms (00:04.679)

postgres=# SELECT AVG(i199) FROM (select i199 from zlayout offset 0) x; -- zedstore
          avg
-----
 500000.500000000000
(1 row)

Time: 379.710 ms
```

# 分区表-大量分区性能提升

• 1000个分区、469倍

- [https://github.com/digoal/blog/blob/master/201905/20190521\\_01.md](https://github.com/digoal/blog/blob/master/201905/20190521_01.md)
- [https://github.com/digoal/blog/blob/master/201903/20190331\\_01.md](https://github.com/digoal/blog/blob/master/201903/20190331_01.md)

case	PG 11	PG 12 beta1	性能提升倍数
单表查询 qps	1161671	1160909	持平
单表upsert qps	317379	332552	持平
分区表(1024分区)查询 qps	1163	545602	469倍
分区表(1024分区)upsert qps	2885	246627	85倍

# GiST index include

- 轨迹，时空搜索
- 按结果集（索引）聚集存储，消除回表IO放大。

# 日志采样

- [https://github.com/digoal/blog/blob/master/201904/20190405\\_09.md](https://github.com/digoal/blog/blob/master/201904/20190405_09.md)
- `log_statement_sample_rate`，当设置了`log_min_duration_statement`时，如果`log_statement_sample_rate`也设置了，它表示百分之多少的超时SQL被记录。
- `log_transaction_sample_rate`，不管其他任何设置，它表示百分之多少的事务被记录duration。（事务为最小粒度单位，如果一个事务被触发了记录duration，这个事务中的所有SQL都会被记录。）

# 目录

- PG 生态
- PG 新版本开发者特性精选
- PG 与 Oracle的全方位对比

从买卖买卖说起

规格参数	
尺寸规格	
室外机净质量	31kg
室内机净质量	10.5kg
内机尺寸	928x300x199mm
外机包装尺寸	920x400x620mm
外机尺寸	860x545x315mm
内机毛重	13.5kg
内机包装尺寸	1000x370x280mm
外机毛重	36kg

技术参数	
制热功率	1300W
制热量	4800W
制冷量	3500W
室内机噪音	41dB
室外机噪音	51dB
制冷功率	800W
电辅加热功率	1050W



## 包装清单

室内机组	1 件
使用安装说明书	1 件
电池	2 件
遥控器	1 件
排水管	1 件
外机机组	1 件
铜连接管	1 件
塑料包扎带	1 件
穿墙管护圈	1 件
腻子	1 件

## 基本参数

空调品牌	AUX/奥克斯
奥克斯空调型号	KFR-35GW/BpR3QYQ1+1
空调功率	大1.5匹
工作方式	变频
空调类型	壁挂式
冷暖类型	冷暖电辅
空调面板颜色	冰玉蓝
适用面积	16-20m <sup>2</sup>
能效等级	一级
智能类型	已连接天猫精灵
上市时间	2018-12

空调	Oracle	PostgreSQL
品牌 (口碑、广告、销量)		国内外用户： 版本发布节奏稳定，产品稳定。
价格		不要钱，类似bsd许可，风险低。 允许云上、线下输出，厂商可以基于PG分发。
核心技术 (压缩机, 变频, 定频, 冷凝剂)	Rac -> pg (polardb, xl, xc, antdb,... 弯道超车) 优化器动态规划 内存表 Tbs(multi datafile-block devs) -> pg(zfs, lvm) Redo group -> pg(zfs, storage) Block level compress	ACID, 单机, 集群(分布式), 多写(pro), 进程, 线程池(13+), bg process, 守护进程, <b>Am(heap,zheap,zedstore,btree,hash,gin,gist,spgist,brin,bloom,rum)</b> Database, schema, tbs, redo, arch, stream, pitr, Compress(block level, openapi), ssl, auth(pwd,ldap,sspi,peer,md5,...) <b>quorum base(多副本),</b> .cbo, <b>geqo(图式规划), aqo, cost const, hint, sr plan</b> Awr, <b>auto parallel(sql, mainantence, idx, bak, restore)</b> HA(shared store, stream rep) <b>Mpp,htap(citus, gpdb, antdb, polardb,...)</b>
功耗 (匹)		Ecs 单机 180万 tpmc total,
能耗比 (功耗, 制冷能力, 制热能力)		104 core , 180万tpmc total => ....
故障率 (品控, 销售量, 故障率)		稳如狗(前提: 基本的部署系统、参数优化)
稳定性 (抗电压波动能力, 变频能力)		Checkpoint动态调度, 波动小。 自动垃圾回收, <b>vacuum freeze</b> 。带调度, 波动小。 大查询资源抢占, 支持用户级并行度配置, 波动小。 锁粒度细, 读写不冲突, 行锁不耗费内存, 没有锁放大问题。
安全性 (防水, 防火, 防短路, 防雷击)		Acl, 认证过程加密, ssl, 敏感信息掩码, tde(13), 存储级加密, 权限控制粒度 ( <b>sec label, advisory lock row, column, table, object, schema, database</b> ) Barrier view

空调	Oracle	PostgreSQL
核心功能 (制冷、制热、除湿。)	Bitmap index -> (pg: gin,rum)	数据类型(普通类型, 数组, json, kv, GIS, 多维, range, comp,uuid,xml,text search,枚举, 字节流, 大对象, 图像, 扩展...), 索引 (btree,hash,gin,gist,spgist,bloom,brin,rum,排他, partial idx, express idx, 虚拟索引, including index) , Fdw(postgres, mysql, oracle,mongo, ms sql, jdbc, odbc, Hadoop, ...) 存储(heap, zheap, zedstore, 未来扩展) Sql语法兼容性, 内置函数、操作符, 自定义函数, 类型, 操作符, 高可用, 在线备份, 增量备份, 增量恢复, 实时容灾, 审计
增值能力: (除甲醛、防霉、预约启动。)	闪回, 闪回query, -> pg(pro snapfs) im-db Datafile direct-io Interval partition 自动内存管理, asm	扩展功能插件数不胜数, gis, 路由, 图, 推荐, 分布式, mm, 流计算, 多维向量, 外部访问源, 逻辑订阅, 资源隔离, HTAP, GPU加速, 并行计算, 传输表空间(pg pro), 块级压缩(zedstore), pg_repack(在线重组) 函数语言扩展 (plpgsql, c, plperl, pljava, pltcl, plpython, pllua, plgo,...)
维护 (自清洁, 可拆洗)		日常维护, 性能监控手段, 优化, 诊断, 上报BUG, 与社区联络非常容易, 可以直接解读错误代码,
产品限制 (打孔, 电流, 安装条件, 使用方面的限制)		支持的芯片, 操作系统, 可以建不限个数db, user, schema, table, 有限取决于block size和编译头文件: column, index, 记录数, 表大小, 数据库大小
售前服务		云、数据库厂商、支持公司、培训公司、高校
售后保障 (安装、保修时长、服务网点分布、价格)		同上(代码级服务除了社区提供, 还有很多公司可以提供)
使用便捷 (即插即用, 遥控, app控制, 智能控制)		语法丰富程度 (最接近sql标准的数据库), gui, cli, 函数开发、调试, develop框架, isv 集成
说明书 (内容丰富程度)		手册, 源码, 书, 搜索引擎, 视频, 培训, 认证, blog, 微信, qq, 钉钉群, 公众号, 5万+

空调	Oracle	PostgreSQL
外观（审美、产品设计）		代码可读性（堪称教科书级），引用各种论文，注释说明丰富。 模块化设计。 开放hook, api，可扩展能力强。
做工（细节处理）		Psql 帮助，tab补齐，快捷命令，快捷帮助、管理函数，框架自动补全，手册index，插件集中化管理。
用料（材料是否环保，有没有偷工减料）		纯C代码。
产能		每年一个大版本，小版本平均1.5个月发布，维护5年。
上市时间		追溯到1973（论文， ingres）， 第一个开源版本始于1995
企业股权分布 注册资本		社区，遍及全球，涉及商业巨头数不胜数 (aws, ms, ibm, google, Alibaba, Tencent, Huawei, apple, .....)
行业认证(3C, ISO) 质量控制。		遵循SQL标准，以及扩展SQL能力。 商业化公司（认证） 社区化运作（管理体系完善）

# 真正的重点！！！

BUG价，还是3折，  
买到就是赚到！！！



# 欢迎交流、PG天天象上



digoal's 微信



**PG技术进阶  
3000+人大群  
每周直播**



关注“阿里数据库技术”公众号

后台回复“PG”获取PPT资料