

大话数据库 终局之战

Digoal

我就是: 偶尔在群里瞎BB的那个德哥!!!

最近BB的作品:

《亿级用户量的实时推荐数据库到底要几毛钱?》

《在数据库中跑全文检索、模糊SQL会不会被开除?》

《刷脸支付会不会刷到别人的钱包?》



江湖名号: 德哥

帮会: PostgreSQL社区

帮会不正当职务: 校长

擅长武器: PostgreSQL

雕虫小技: 全文检索、模糊搜索、化学分析、
人脸识别、相似推荐、时空调度、向量搜索...

必杀技: 删库跑路

正经职务: 阿里云数据库产品经理

愿景: **没有MyBase解决不了的问题,如果有,就多买几台**

格言: 公益是一辈子的事

业余爱好: 写博、写专利、分享、毁人不倦

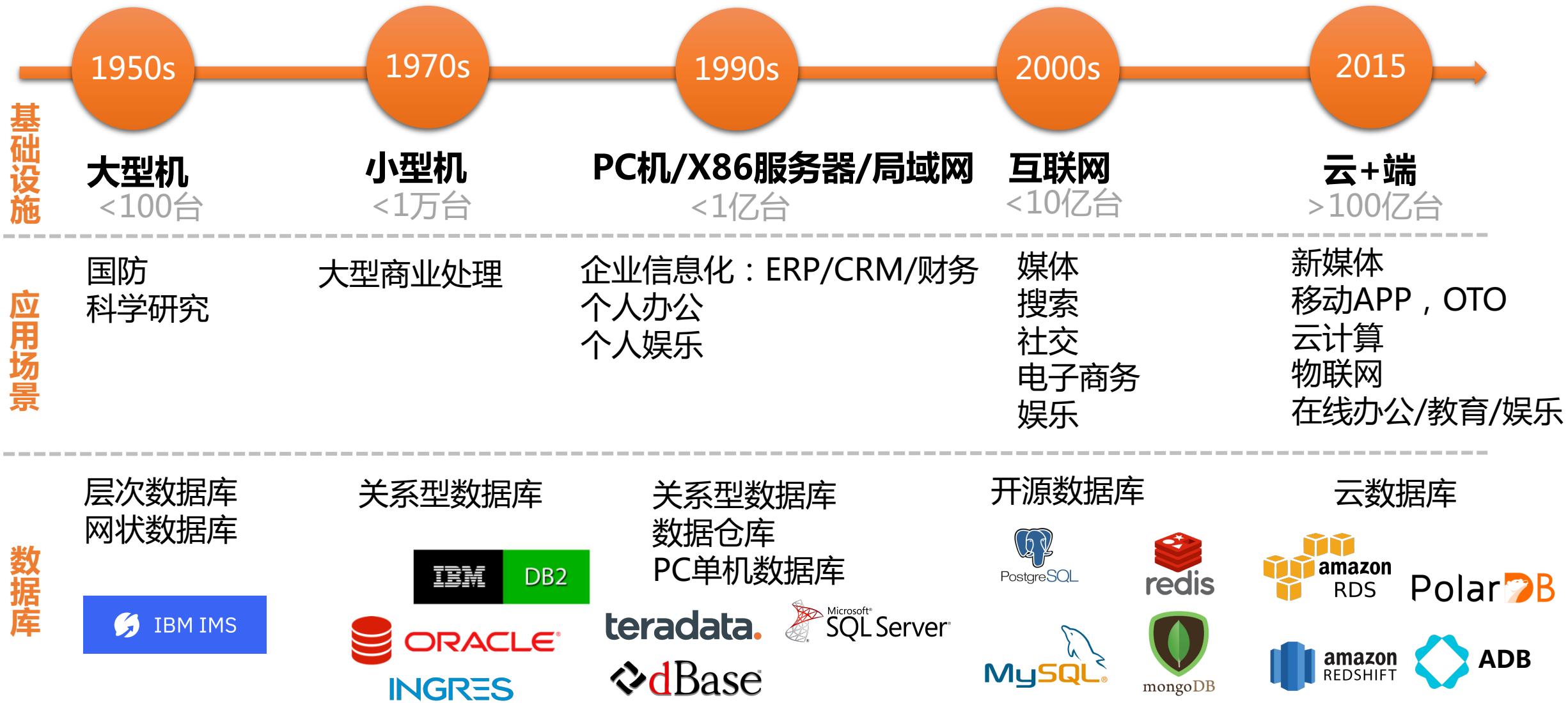
github: <https://github.com/digoal>

联系方式:

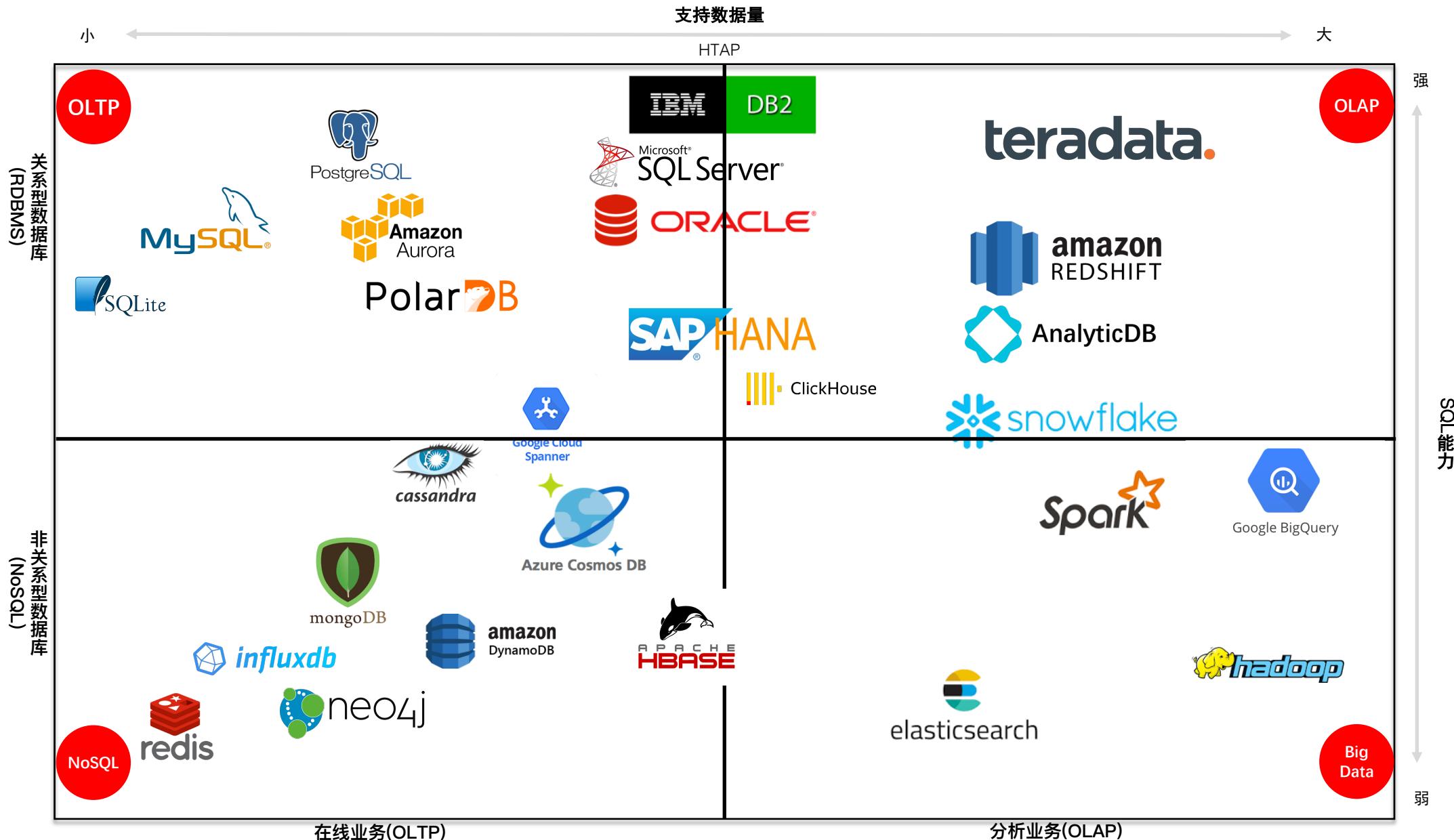
目录

- 数据库发展历史
- 第一次数据库大战
- 第二次数据库大战
- 第三次数据库大战
- 第四次数据库大战
- 大话云数据库的变革 - 下一代云数据库形态

数据库发展历程



数据库分类



第一次数据库大战(1990s)

- 开放系统 VS 封闭系统

x86(80年代, 90年代已成熟)

Windows, Linux

SQL Server, Oracle, Sybase,
Infomix

开放体系, 便宜、人才多

IBM 大、小机

AIX

DB2

封闭体系, 贵、人才少

- 导火索
 - 80年代, 开放式硬件x86兴起,
 - 90年代, 开放式硬件操作系统Windows, Linux兴起,
- 护城河
 - Oracle: 大举收购应用软件(例如people soft).
 - SQL Server: Windows+应用软件.
 - sybase: 无护城河, 被软件厂商抛弃. 最终卖给SAP(办公, erp软件行业霸主).
 - infomix: 无护城河, 被IBM收购, 后有"卖给"南大通用.
- 衰败分析
 - DB2, 封闭体系, 贵、人才少, 缺乏增量应用, 死循环.

第二次数据库大战(2000s)

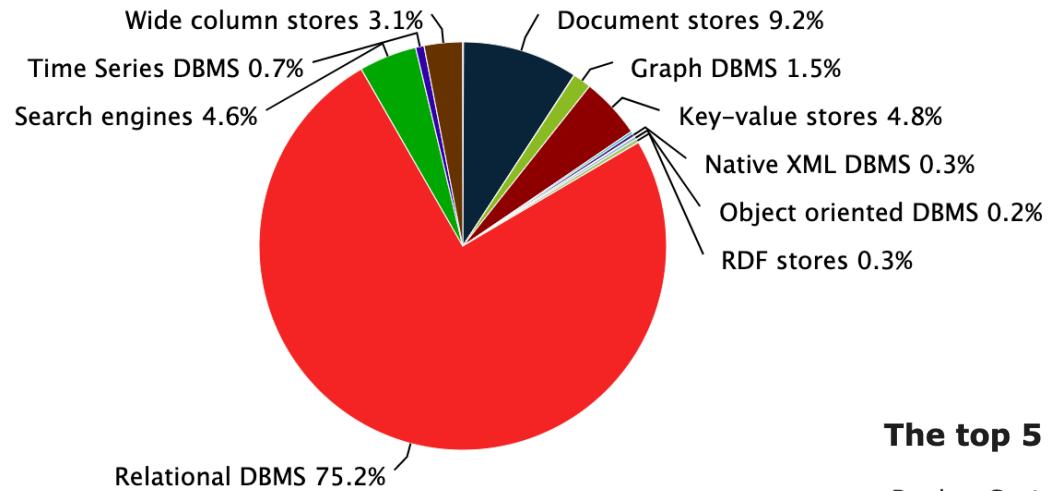
- 开源 VS 商业

MySQL, PostgreSQL, Mongo
其他开源产品不探讨(Hadoop等)

Oracle, SQL Server

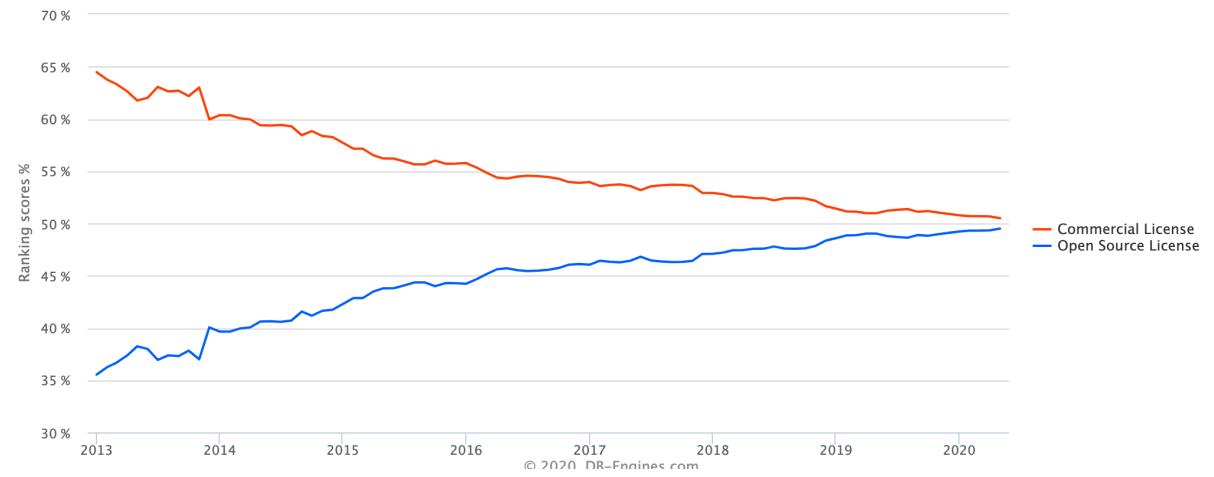
DB-Engines(357种数据库 , 2020.05)

关系型 & NoSQL



商业 & 开源生态

Popularity trend



The top 5 commercial systems, May 2020

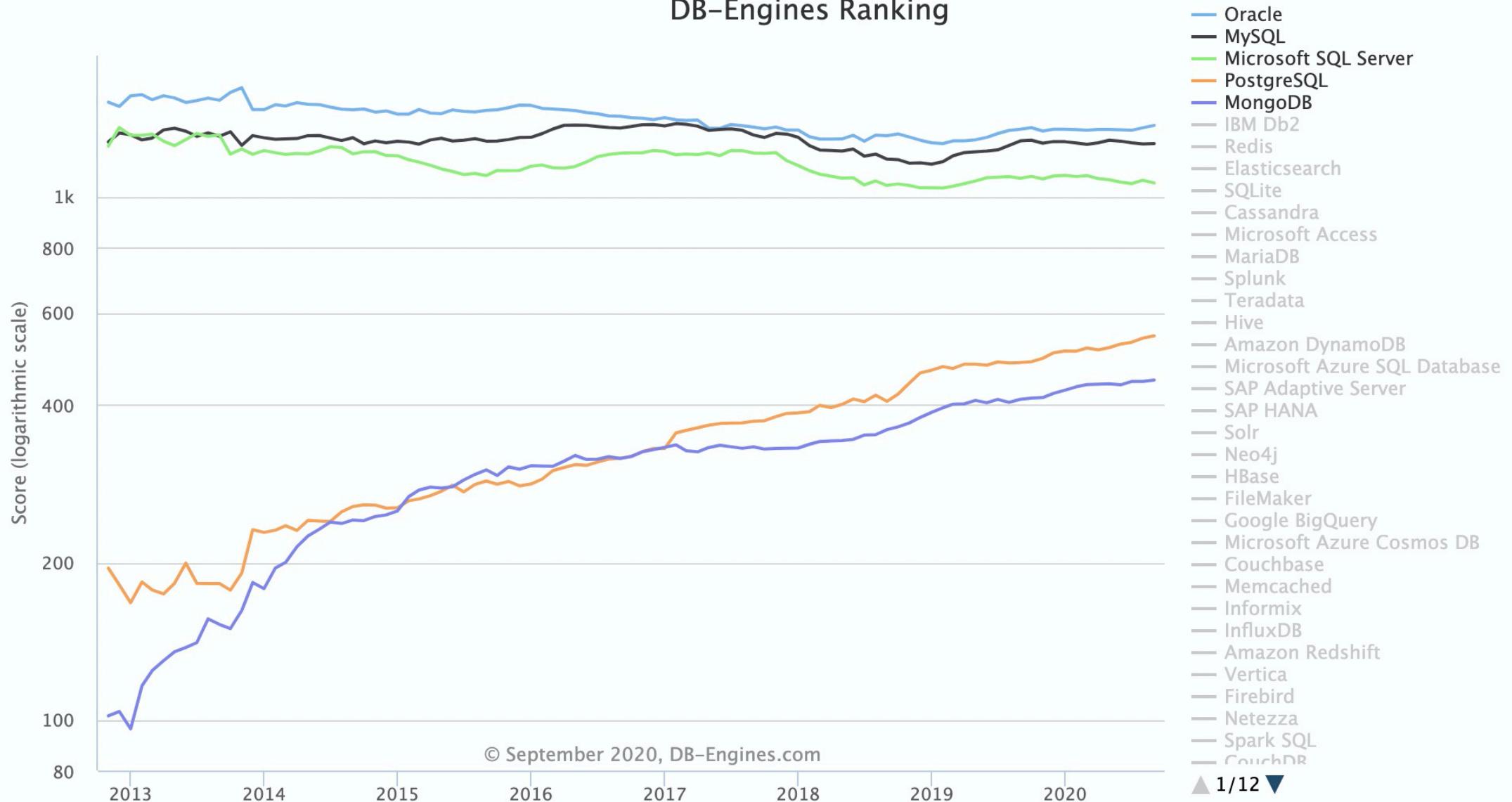
Rank	System	Score	Overall Rank
1.	Oracle	1345	1.
2.	Microsoft SQL Server	1078	3.
3.	IBM Db2	163	6.
4.	Microsoft Access	120	10.
5.	Splunk	88	13.

The top 5 open source systems, May 2020

Rank	System	Score	Overall Rank
1.	MySQL	1283	2.
2.	PostgreSQL	515	4.
3.	MongoDB	439	5.
4.	Elasticsearch	149	7.
5.	Redis	143	8.

- MySQL
 - 转折-互联网兴起,个人站点爆发
 - 护城河-轻量(简单易用够用 流行场景覆盖广)人多
- Mongo
 - 转折-敏捷开发(schema less)兴起
 - 护城河-超级轻量(简单易用够用 流行场景覆盖广)人多
- PostgreSQL
 - 转折-企业软件PaaS化 (终身买断(贵) ~ 订阅(便宜), 企业软件开发商必须降成本(商业数据库遭殃))
 - 护城河-类BSD开源许可(**基于PG的商业化门槛低**),可编程OPENAPI(**代码可塑门槛低**), 稳定可靠、功能对齐(**应用开发门槛低**).
 - 大量基于PG的衍生商业数据库.
 - 大型企业复杂业务选择PG.
 - 商业数据库低成本替代方案.

DB-Engines Ranking



第三次数据库大战(2010s)

- 云 VS 开源
- 导火索
 - IDC 没落, 云厂商崛起, 云OS诞生.
- 云数据库DBaaS服务
 - 攻擂策略: 标准化产品.
 - 更快、更稳、更安全、更智能.
 - 开箱即用.
- 商业开源数据库产品
 - 防守策略: 开源协议.



飞天 (阿里云自主研发计算机操作系统)

 编辑

 讨论

 上传视频

飞天 (Apsara) 是由阿里云自主研发、服务全球的超大规模通用计算[操作系统](#)。

它可以将遍布全球的百万级服务器连成一台[超级计算机](#)，以在线公共服务的方式为社会提供[计算能力](#)。

从PC互联网到移动互联网到万物[互联网](#)，互联网成为世界新的基础设施。飞天希望解决人类计算的规模、效率和安全问题。
飞天的革命性在于将[云计算](#)的三个方向整合起来：提供足够强大的计算能力，提供通用的计算能力，提供普惠的计算能力。

飞天诞生于2009年2月，为全球200多个国家和地区的创新创业企业、政府、机构等提供服务。

软件名称	飞天操作系统	软件语言	Apsara
开发商	阿里云	诞 生	2009年2月

第四次数据库大战(2020s)

- 云 VS 开源 VS 商业

三足鼎立

第一、二次大战:
商业数据库时代用户:
渠道驱动(传统行业)

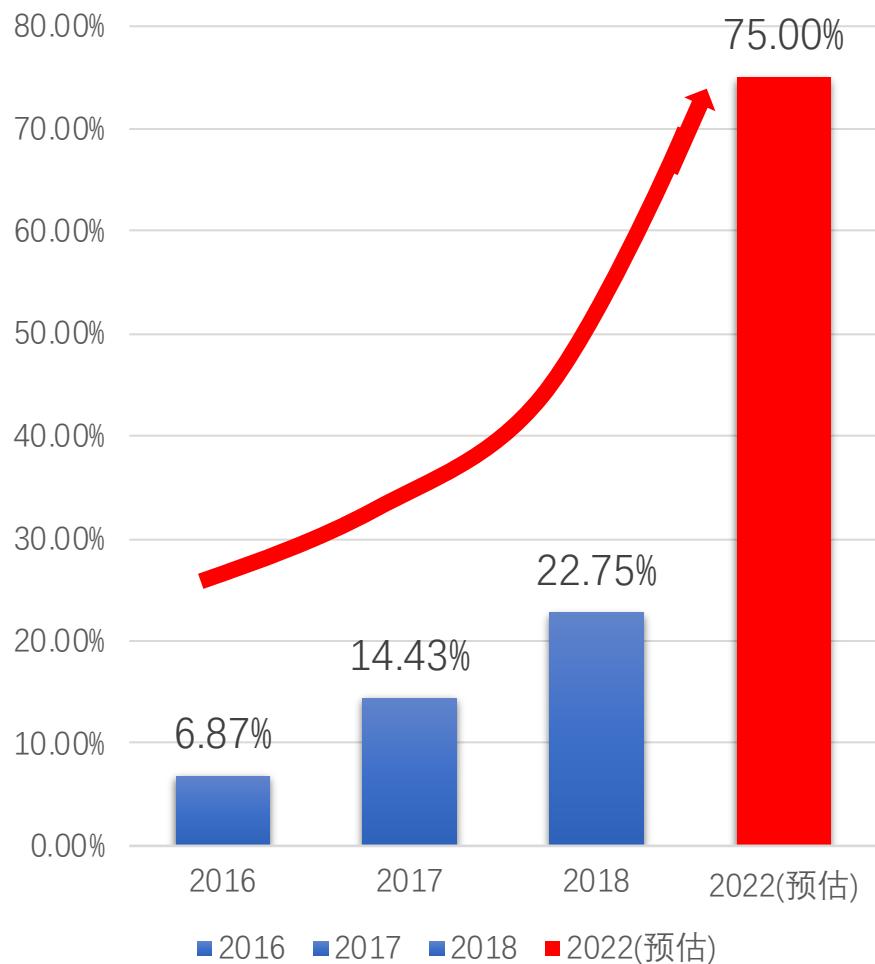
第三次大战:
开源数据库时代用户:
技术驱动(互联网颠覆传统行业)

用户也在变

第四次大战:
数据库服务时代用户:
业务驱动(技术优势被抹平, 业务驱动)

某短视频、直播、电商网站:
全站基于阿里云, 50个研发,
6个月用户突破1亿, 日UV达
3500万.

云数据库快速发展，(Gartner报告)2022预计75%的数据库都在云上



魔力象限中云数据库厂商成长最快

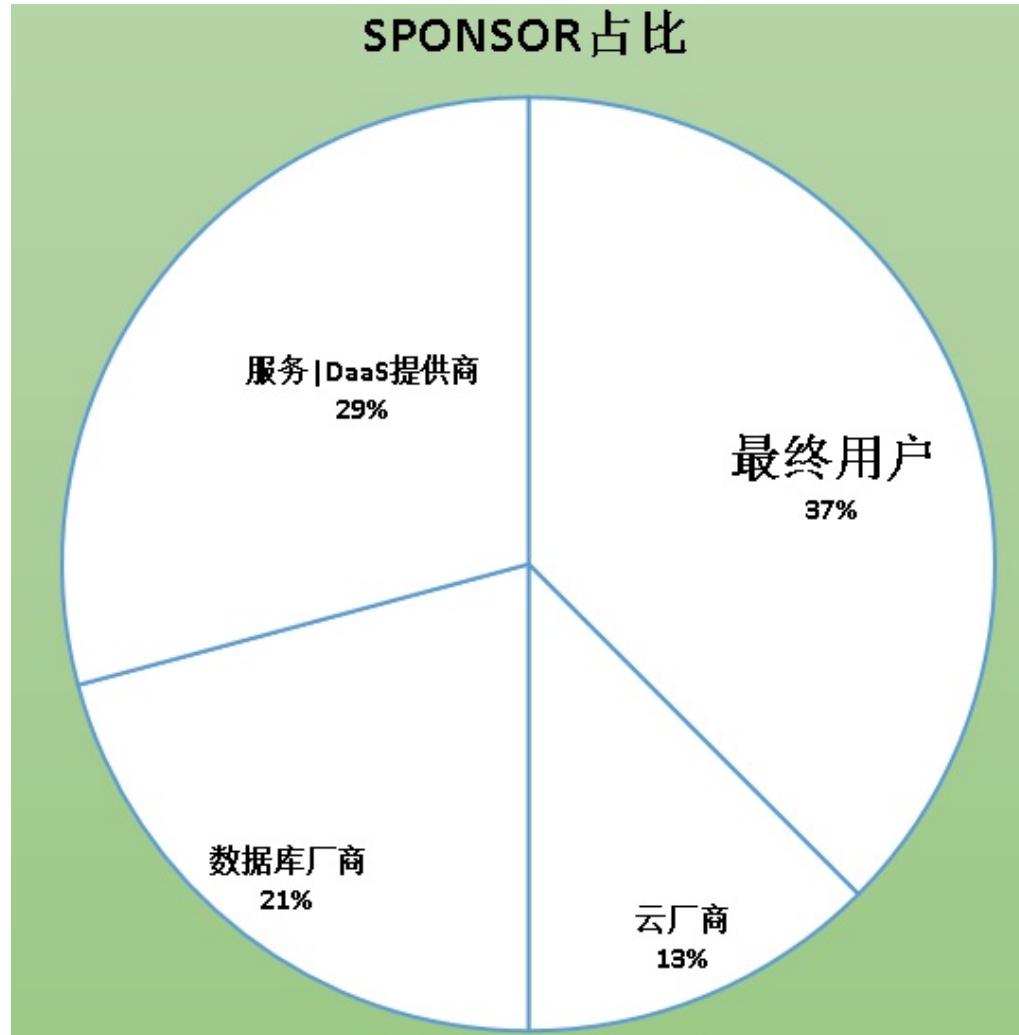


© Gartner, Inc

大话终局之战 – 三足鼎立

- 导火索
 - ALL in Cloud
- 商业开源数据库,
 - 既要求发展, 又要防麻雀.
 - 拓新: 差异化市场. 避开与巨头正面杠.
 - 守城:
 - 开源生态, 培养开发者习惯、绑定应用护城河.
 - 防给他人做嫁衣, 协议约束.
 - 积极与云商业合作, 避免正面交锋.
- 纯社区开源数据库,
 - 背后的核心贡献者利益集团. PG为例(用户、支持商、厂商、云)

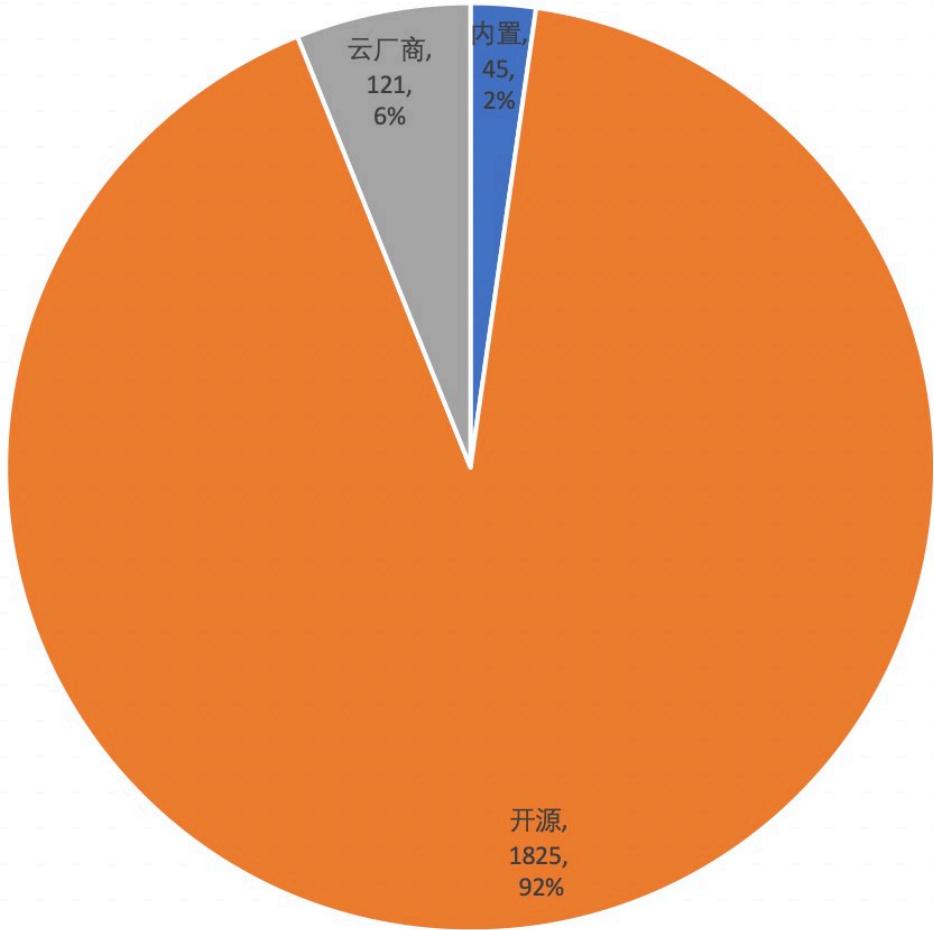
PG 为 例，社 区 开 源 数据 库 !



- 用户
 - 以小博大, 1~2个内核研发长期投入, 博大整个社区, 可以持续用上成熟稳定、安全的产品.
- 支持商
 - 市场? 本想把市场培养起来卖服务(培训、支持), 但是随着数据库越来越智能, 云厂商的智能化服务产品出现. 服务市场萎缩, 怎么办?
 - 引导PG数据库架构变革, 转型行业应用开发商.
 - 开放更多应用领域openapi接口, 让数据库变成应用数据处理平台, 而不是与应用脱节的数据库. 弱化数据库概念, 增强应用概念. 只有这样支持商的市场盘才能做大, 同时不与云厂商起争端, 同时又能满足用户诉求, 继续做大盘子.
- 厂商
 - 和谁对抗? 云? (100kg vs 1000000kg级)?
 - 依托开源生态, 避免重复造轮子.
 - 避开锋芒, 主攻非云、非充分竞争市场.
- 云
 - DBaaS产品, 吃开源红利.
 - 基于开源自研, 贡献提高影响力, 避免重复造轮子, 做好兼容, 继续吃开源红利. 对抗纯商业数据库市场.

开放应用层OPENAPI的数据库长啥样

PG 插件数量分布



开放应用层OPENAPI的数据库长啥样

- 图像识别 - 阿里云PG PASE模块
- 地理信息 – 阿里云PG Ganos模块
- 模糊搜索 – 阿里云PG 中文分词, 模糊查询加速模块
- 化学分析 – 阿里云PG rdkit 模块
- 实时推荐 – 阿里云PG hll 模块
- 精准营销 – 阿里云PG roaringbitmap 模块
- 流处理 – piplinedb
- 机器学习 – madlib
- 图计算 - agensgraph ext
-

大话终局之战 – 三足鼎立

- 商业数据库
 - 拓新: 收购为主. 收割盗版为辅.
 - 用户为什么没有逃离? - 应用.
 - 目前仅存的两款流行商业数据库sql server, oracle 背后都是应用在支撑. 应用在, 数据库就在. 让用户换应用更难, 大多数软件都支持sql server, oracle. 等于要颠覆整个软件行业(正在发生).
 - 守城: 应用上云不可逆转, 积极开展自建云, 混合云, 私有云业务.
 - 继续拓展自己的应用地盘, 让用户上应用云, 用户应用和数据都上了云, 就更难逃离. 例如微软, oracle.
 - 云数据库
 - 拓新:
 - 比自建、三方 dbaas over cloud 更优 vs 开源
 - 自研 vs 商业(包括开源企业版)
 - 商业开源数据库很诱人, 如mysql, 但是养大它等于养大对手. 表象上: 阶段性吸收红利行为.
 - 社区开源数据库, 开源许可允许你改、卖、有自主产权. 重兵投入, 砸出一个未来.
 - 布局云生态和应用生态, 应用为王, 让更多应用支持自己的云原生数据库.
 - 守城:
 - 云智能生态, 与云上其他应用紧密结合形成整体护城河.

大话云数据库的变革： 下一代云数据库形态

• MyBase 专属集群发展的时代背景

All in Cloud, 用户提出更高要求:

- 既要: 稳定、高效、便捷
- 又要: 性价比、安全可控
- 还要: 可定制化功能、出任何问题有厂商兜底

阿里云数据库发展历程

<https://www.aliyun.com/product/apsaradb/cddc>

2012



RDS
云托管

2015



NoSQL
云托管

2017



PolarDB
云原生

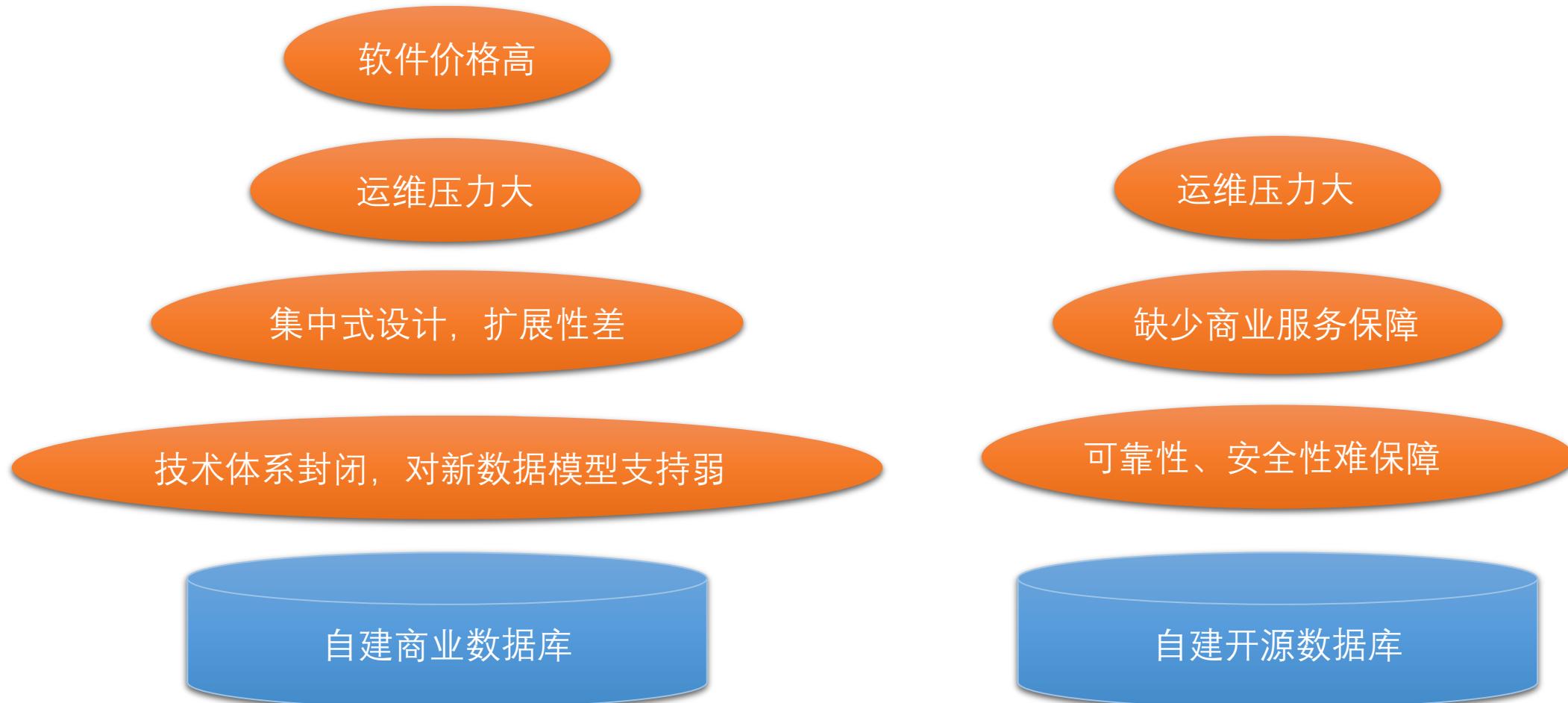
2019



MyBase
云专属

一种新的面向中大客户的形态
支持MySQL/SQL Server/PG/Redis/MongoDB/HBase

传统自建数据库的挑战



阿里云数据库



PolarDB



10万客户
40万实例

- 中国最强的数据库研发团队
- 提供了最丰富的数据库选型，PostgreSQL、MySQL、SQLServer、Redis、Mongo等
- 100%兼容生态，客户业务无需修改，可以直接上云
- 所有的内核都是经过了双十一高并发与稳定性考验

开箱即用

更快

更稳

更安全

- 1分钟创建实例
- 自动参数调优配置
- 内置监控告警
- 一键数据迁移

- 内核优化
- 参数优化
- 性能至少提升30%
- 快速弹性扩展

- 可用性SLA：99.99%(业界第一)
- 稳定硬件选型与压测
- 管理了数十万实例的HA系统
- 快速内核bug修复

- 内置备份恢复，避免删库跑路
- 可以恢复到一周内任意时间点
- 回收站，避免误操作
- SQL/操作审计
- 传输与存储加密：TDE、SSL
- PCI DSS/ISO/SOC/等保安全认证

中大客户对数据库更高要求

全托管云数据库 服务解决了	还需要什么？
高可用	综合成本 (购买成本、迁移成本、系统利旧，维护成本)
安全可靠	稳定可控 (稳定安全、权限开放、无缝上云、DBA运维能力保留)
高性能	解决规模瓶颈 (大容量，高并发、高性能)
强扩展	帮助业务转型 (完整服务、解决方案、代差能力)

MyBase：面向中大客户自主可控、极致性价比的专属形态

资源独享

成本可控

灵活调度
弹性伸缩

安全可控
可自运维

阿里内核

完整方案

用户级别资源隔离，
提供主机售卖，完全的
资源隔离，避免因
资源争抢导致的业务
不稳定

更低成本，通过合理的
资源超卖，实现更高的
资源利用率降低
成本

提供紧凑型，均衡性
资源分配策略，同时
提供灵活的备库重搭，
实例迁移，主机替换
等能力，支持指定主
机部署实例。

可以在特定时间对特
定实例进行临时的资
源弹性扩展，满足暴
涨需求

开放 OS 系统，提供
最大的自主性和灵活
性，支持定制化脚本
安装，满足中大型企
业定制化数据库服务
需求

企业级数据库内核，
对开源数据内核进行
了大量增强和优化

全托管数据库服务，
提供安装，部署，高
可用，监控，告警，
账号管理，自动优化，
备份还原等全套数据
库服务能力

一种新的面向中大客户的形态

支持MySQL/SQL Server/PG/Redis/MongoDB/HBase

憋废话了，到底哪里有 MyBase？

- 入钉钉群，**咨询砖家**
- **免费** 活动进行中!!!
- **周周** 有直播和红包雨!!!

