

创新技术推动应用进步

# 库外计算提升数据库应用品质

 数据分析师  
[www.cda.cn](http://www.cda.cn)

**RAQSOFT**



## 库外计算引擎

润乾集算器



## 减轻数据库负担

**CDA** 数据分析师  
www.cda.cn  
存储过程、冗余中间表、多样性数据



## 提高数据库性能

执行路径、并行取数、跨库集群



## 浅谈数据中心

# 目录



# 库外计算引擎



## 润乾集算器

面向（半）结构化数据计算的程序设计语言  
提供丰富的计算类库，支持多线程并行

数据库外的计算引擎



## 库外计算

计算是数据库的重要任务

数据库的封闭性



## 存储过程的目的

数据整理

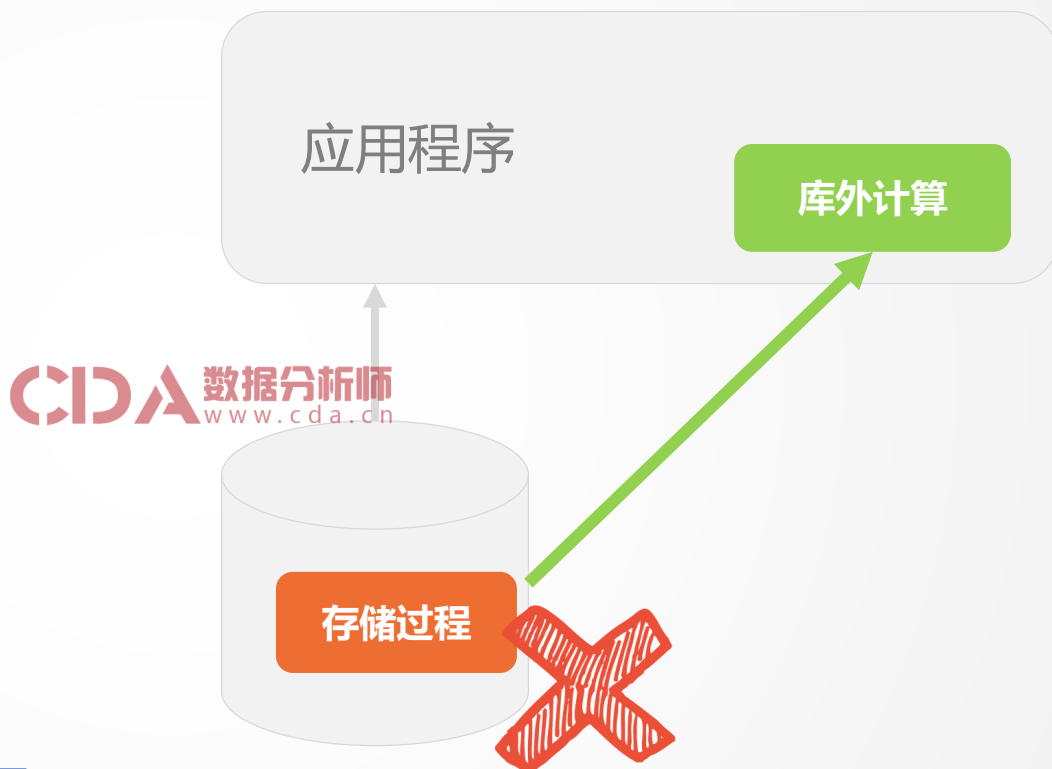
报表查询

## 存储过程的问题

应用内与应用间耦合

安全性与易管理性

## 库外计算替代存储过程





## 中间表的由来

运算复杂或数据量大

再次计算的能力

## 中间表的问题

数量众多占用数据库资源

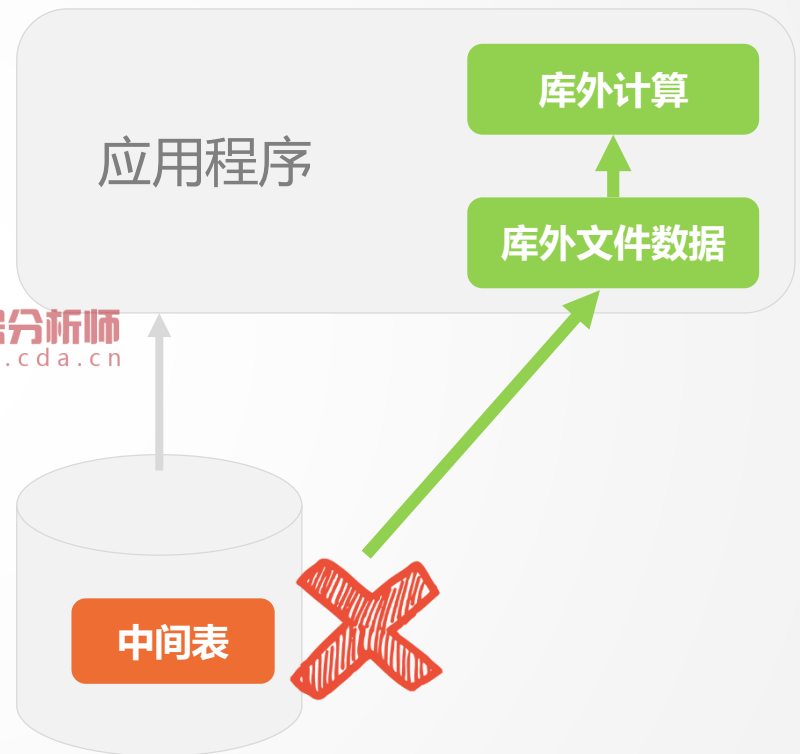
线性结构导致管理困难

## 库外计算将中间数据外置

计算不依赖于数据库，其它能力不需要

绑定应用、树状结构、易于管理

CDA 数据分析师  
www.cda.cn





## 多样性数据普遍存在

txt, csv, xlsx

nosql, hadoop

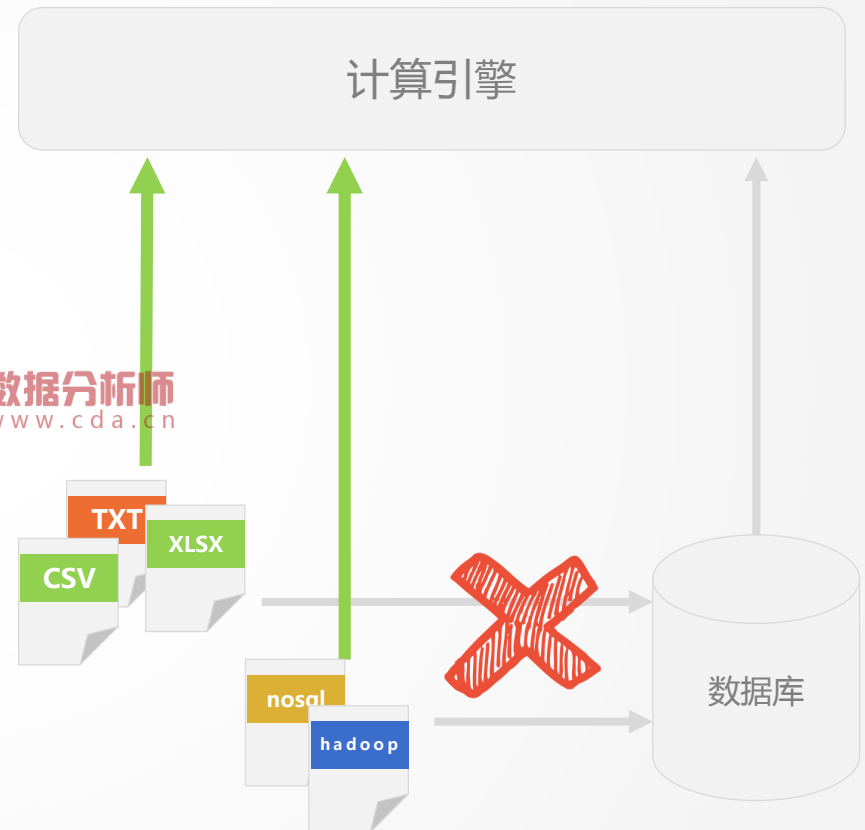
json, xml

## 直接计算多样性数据

不需要建设专门的数据库及转入工作

结构简单、实时性更好

CDA 数据分析师  
www.cda.cn





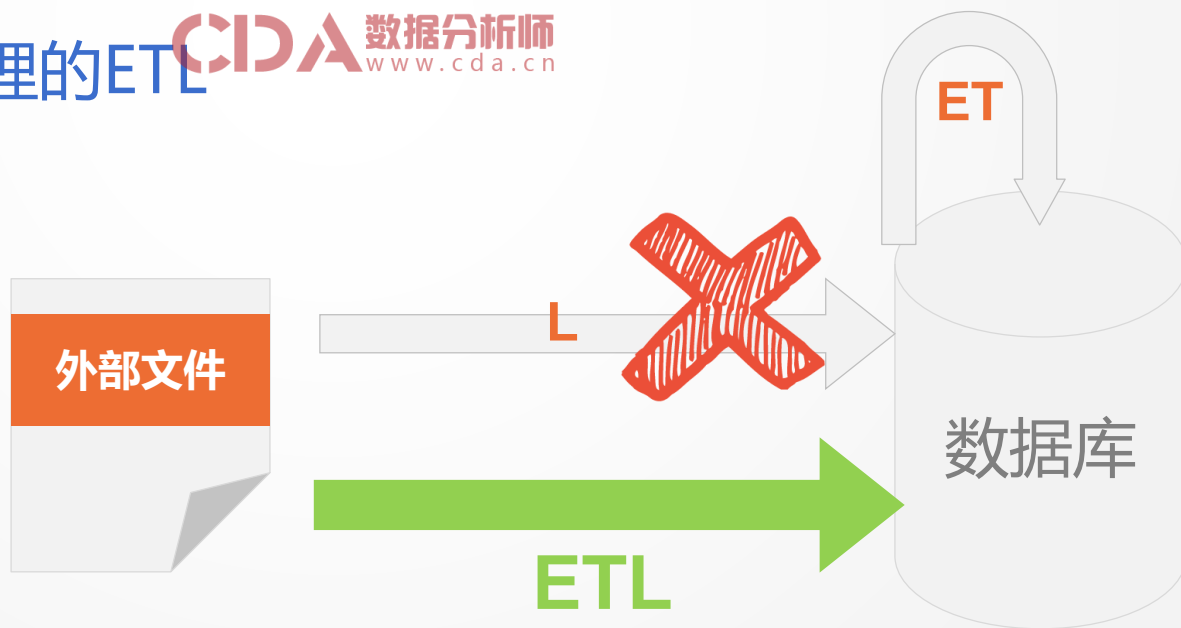
## 缺乏库外计算能力扰乱ETL过程

ETL ? ELT ? LET ?

加大数据库负担

库外计算实现合理的ETL

CDA 数据分析师  
www.cda.cn





复杂SQL的执行路径难以控制

库外计算优化SQL执行路径

自由控制执行步骤

部分运算移至库外进行

CDA 数据分析师  
www.cda.cn







# 并行取数



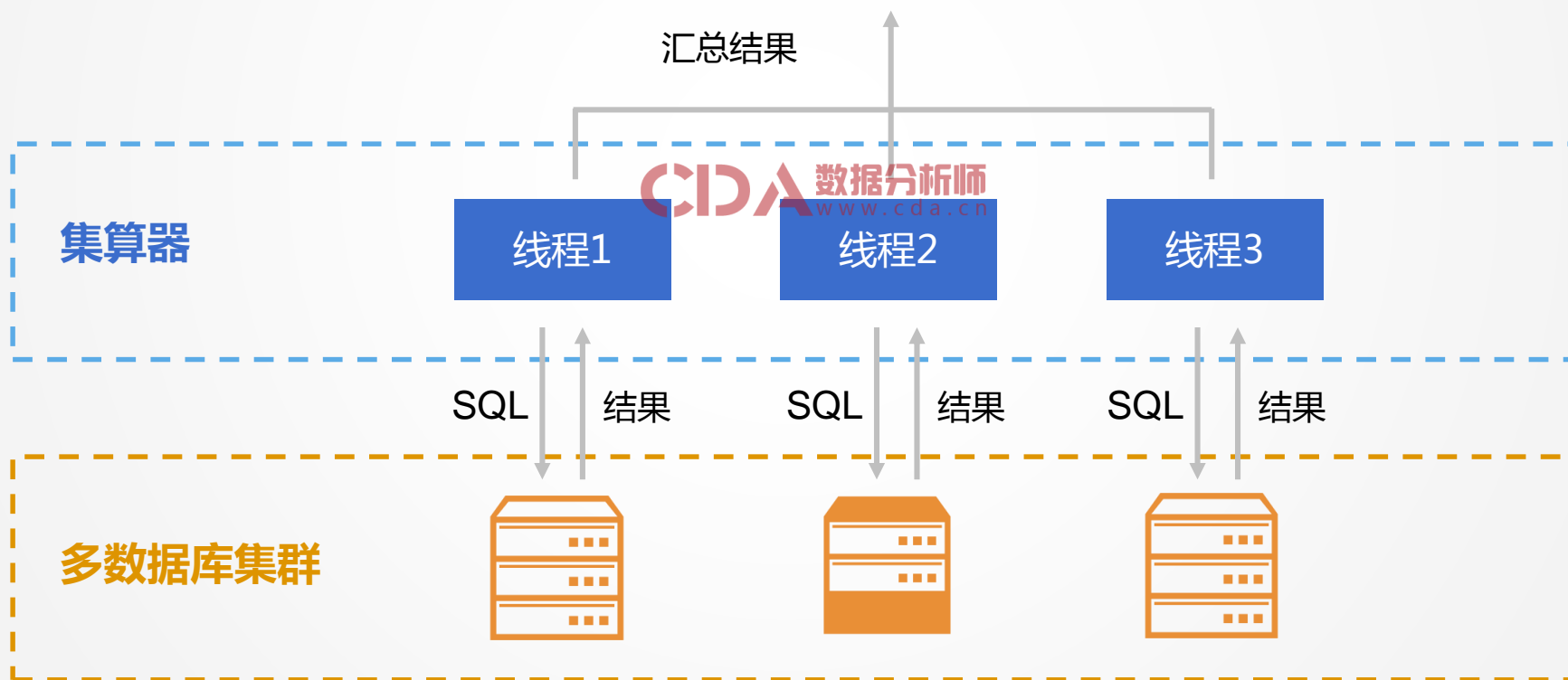
## JDBC性能瓶颈

## 计算引擎多线程取数

	A	B	C
1	fork 4	=connect(db)	/分4线程
2		=B1.query@x("select * from T where part=?",A1)	/分别取每一段
3	=A1.conj()		/合并结果



## 异构数据库集群





# T+0查询报表



## T+0问题

交易一致性要求关系数据库

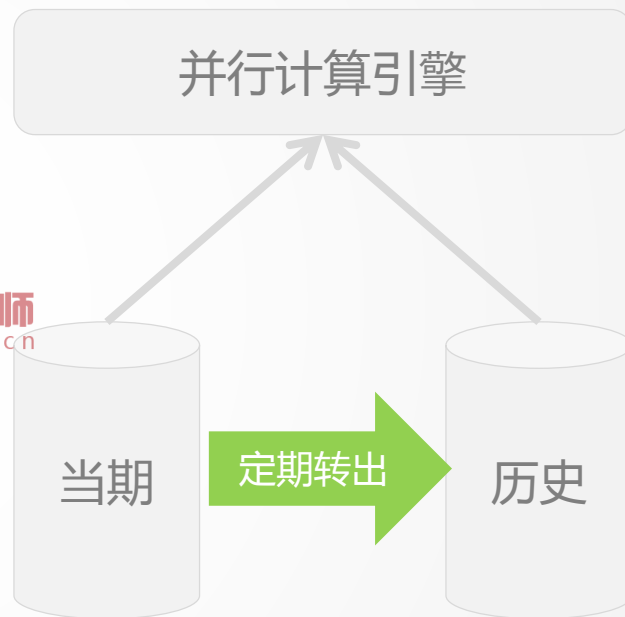
历史与当期同库，数据量太大

历史与当期异库，跨库计算困难

CDA 数据分析师  
www.cda.cn

## 库外计算实现并行跨库计算

历史数据还可文件化





## 数据中心的特征与要求

数据库群

多样性数据源

服务式接口

访问受控

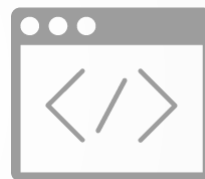
数据脱敏

## 库外计算引擎实现数据库中心访问层

CDA 数据分析师  
www.cda.cn



计算能力



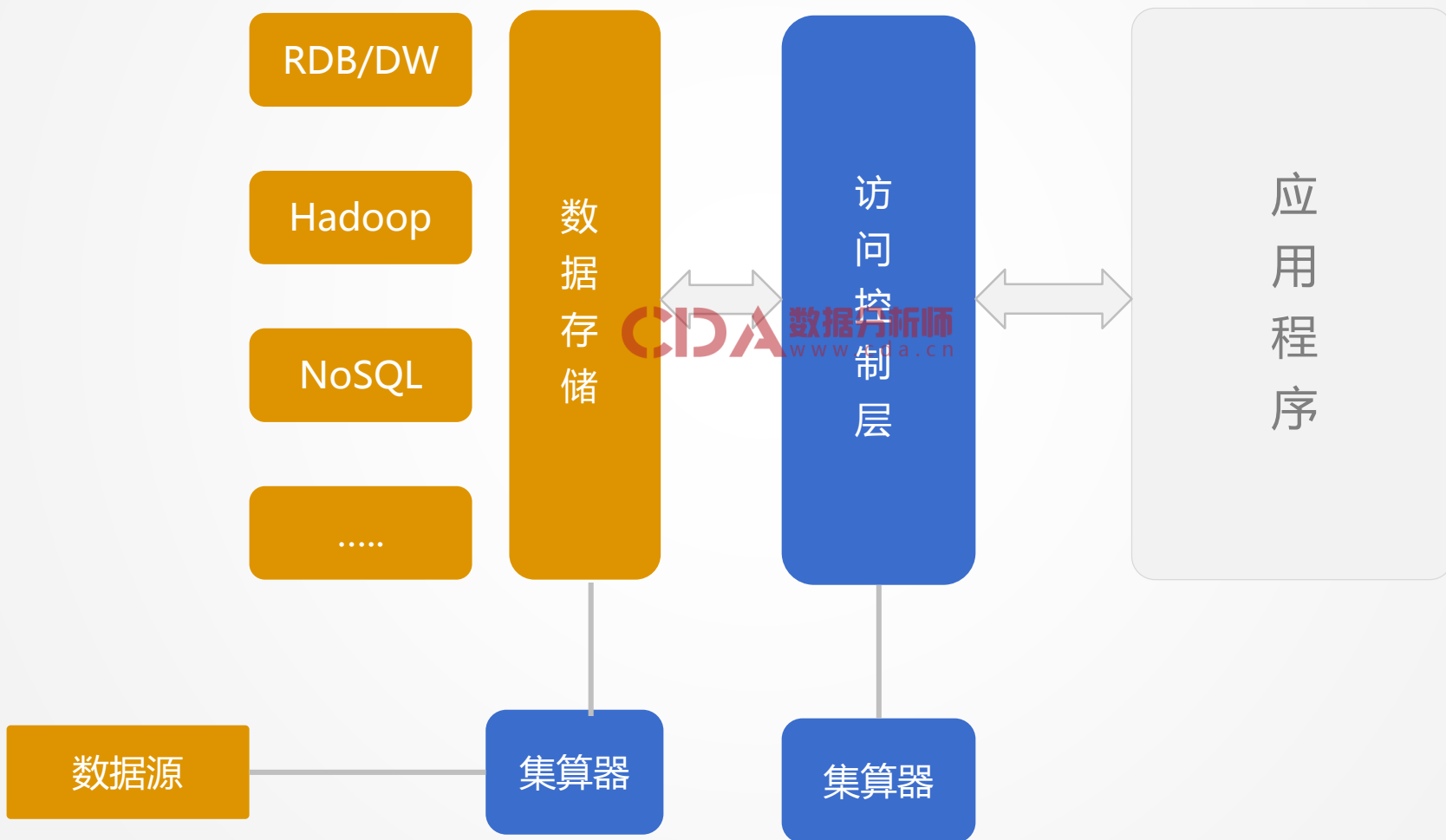
编程能力



集成能力



# 数据中心结构





润乾

RAQSOFT

创新技术 推动应用进步

[www.raqsoft.com.cn](http://www.raqsoft.com.cn)

CDA 数据分析师  
[www.cda.cn](http://www.cda.cn)

谢谢大家！