



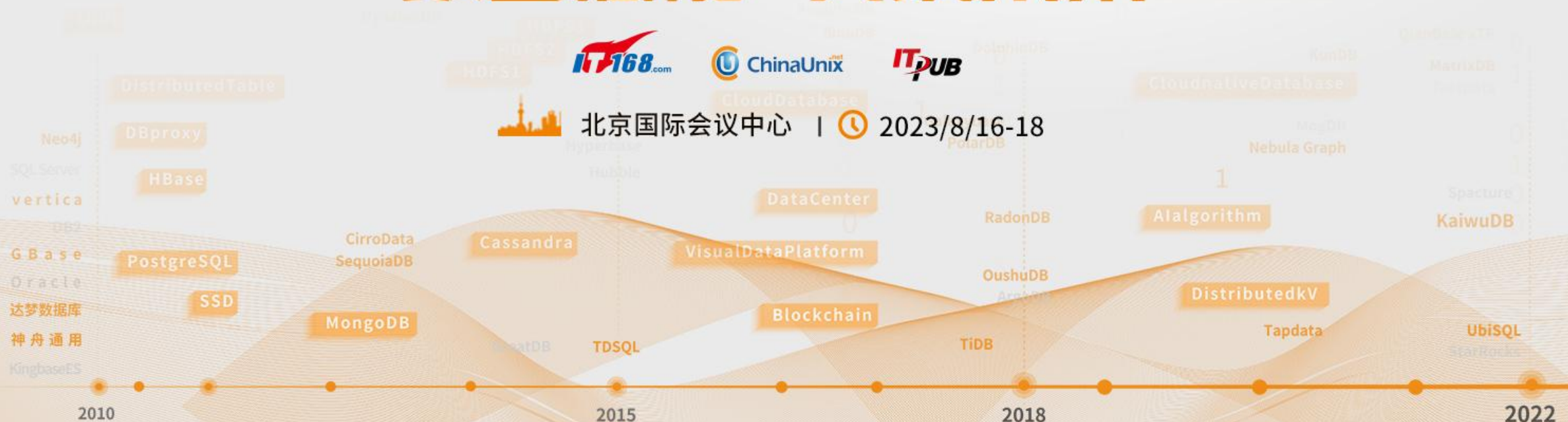
第十四届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA

数智赋能 共筑未来



北京国际会议中心 | 2023/8/16-18



本来生活网的大数据演变

本来生活网 DBA经理 姚伟

目录
CATALOGUE

0

公司概况

0

本来生活网大数据平台建设历史

0

本来生活网大数据平台现状

0

本来生活网大数据平台展望

Part 1 公司概况

本来生活网介绍

Part 1 公司概况

本来生活网：始终致力于改善食品安全现状、倡导健康生活方式

本来生活网创立于2012年，是中国农产品品牌化的领导者，优质生鲜DTC电商平台。

从国内优质食品供应基地、国外优质食品供应商中精挑细选，剔除中间环节，提供冷链配送、食材食品直送到家服务。

我们提供的产品包括：健康安全的蔬菜水果、肉禽蛋奶、米面粮油、母婴童、熟食面点等；具有深厚积淀和历史传承的优质原产地食品；拥有品质保障的进口食品等。

本来生活网独创B2C+DTC模式，以全球买手直采、特色食品定制、自有品牌开发等为业务特色，建立起有独特价值的生鲜供应链体系，赋能包括褚橙、李玉双大米、俞三男状元蟹等400多个生鲜农产品品牌。

本来生活网拥有自有电商平台渠道、第三方平台渠道、私域流量渠道和企业团购渠道等，实现全渠道覆盖；为消费者提供涵盖水果、蔬菜、肉禽、水产、熟食、粮油、乳品母婴、酒水、厨具、日用百货、休闲食品、鲜花、保健品、宠物食品及用品、软百货、3C数码、文体玩具等17大品类、24,000多款优质商品。目前，本来生活网有优质用户3,600万，服务30万企业客户。

2020年12月起，本来生活网实现生鲜冷链直达配送服务覆盖全国31个省级行政区。

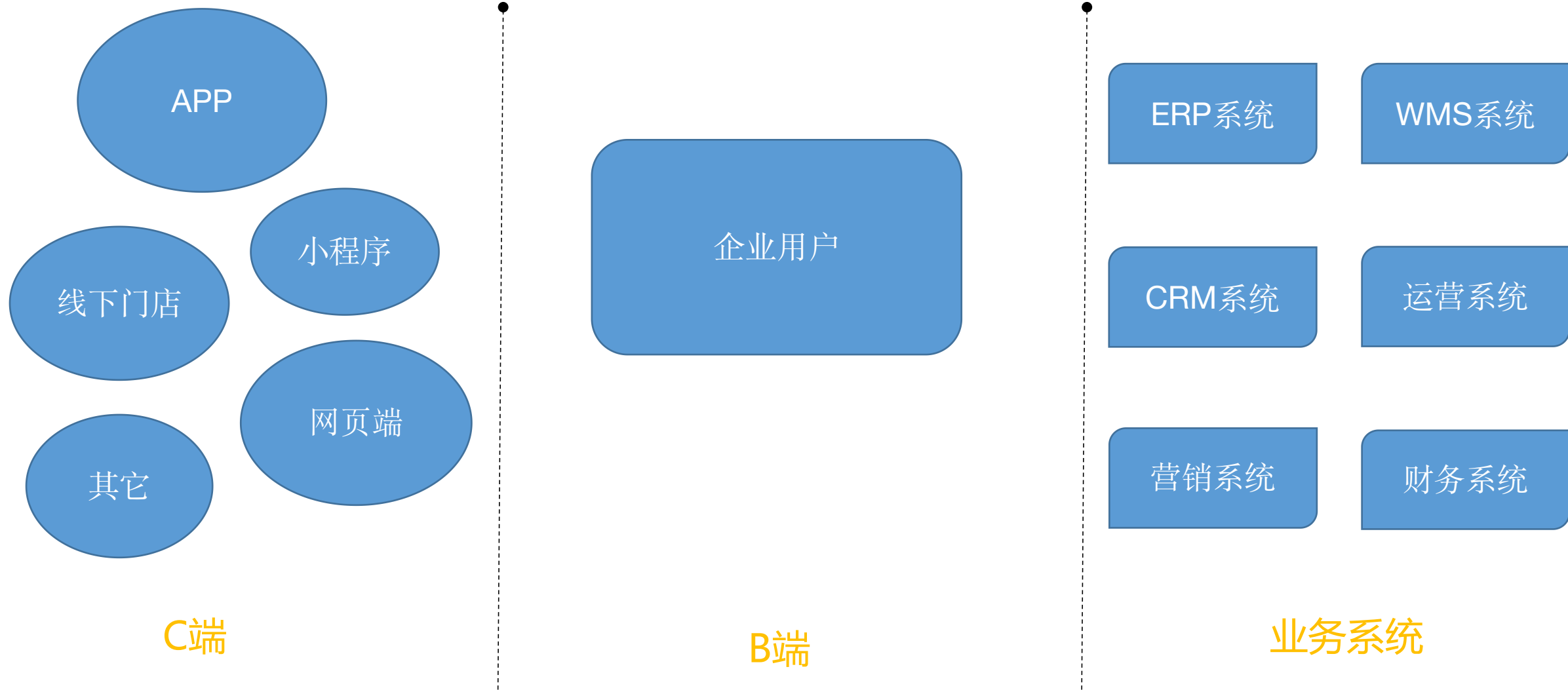
2018年至2021年连续四年全财年盈利。

Part 2

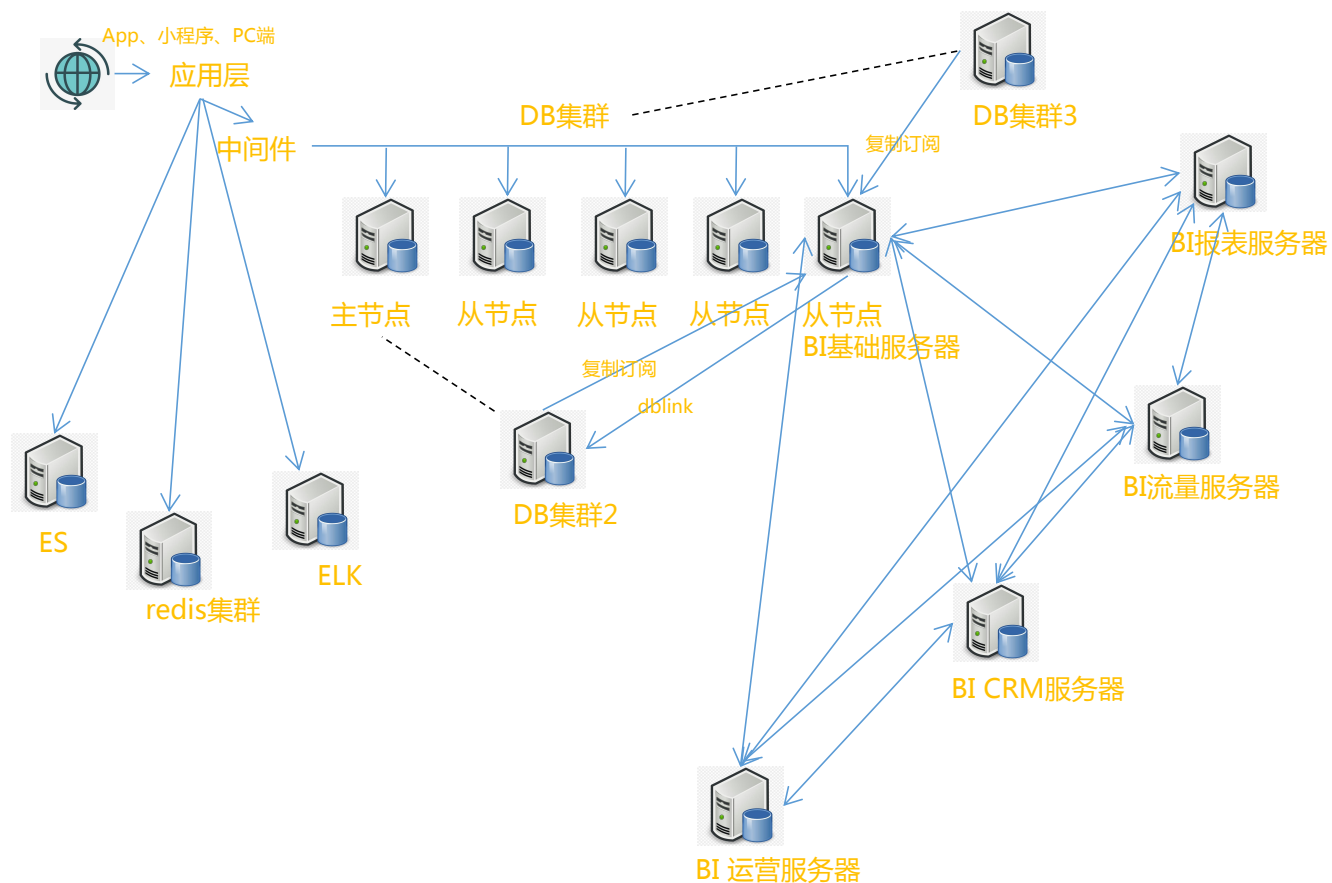
本来生活网大数据平台建设历史

本来生活网成立之初的大数据建设

本来生活网业务简介



早期大数据服务架构

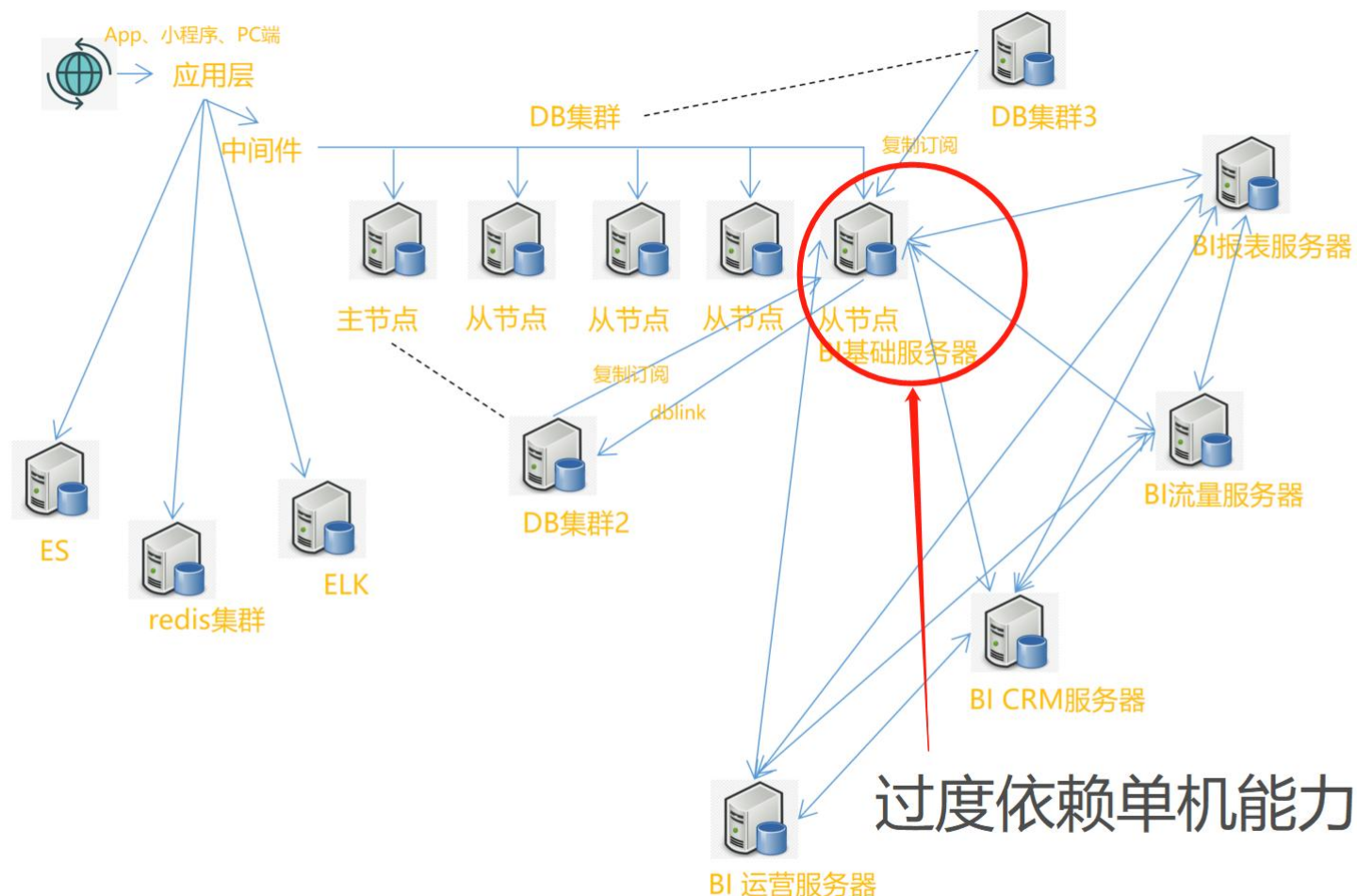


痛点

1. 单机处理数据有极限，数据量一大，出数据变慢
2. 数据量一大，单机压力剧增，实时数据报表很难保证
3. 用户需求越来越多样化，越来越多的定制化需求
4. 服务架构复杂，一个环节出问题就可能导致报表出不了
5. 扩展性差，遇到性能瓶颈只能依靠降低数据量和换机器

- 销售报表
- CRM报表
- 月结报表
- 运营报表
- 财务报表
- 定制报表
- 营收报表
- 仓库报表

以前的数据实时化实现



问题

1. 因为所有数据都会通过主从同步、复制订阅和dblink实时汇总数据到BI基础服务器，然后根据SQL语句去实时取数，一旦单机压力过大就会导致实时取数很慢
2. 因为单机性能有限，只能数次去提升单机的硬件性能，但是终有上限
3. 扩展性差，因为用数据库实时出数据，最多只能对BI数据库做主从，但是同时会增加复制订阅和dblink的设置，导致服务架构更加复杂和不稳定

历史实时数据

数据库集群的备机作为BI数据源

其它数据通过复制订阅实时同步、dblink跨服务器查询获得

对单机性能要求高

数据量一大会导致性能瓶颈，尤其是内存不足

整体服务架构过于复杂

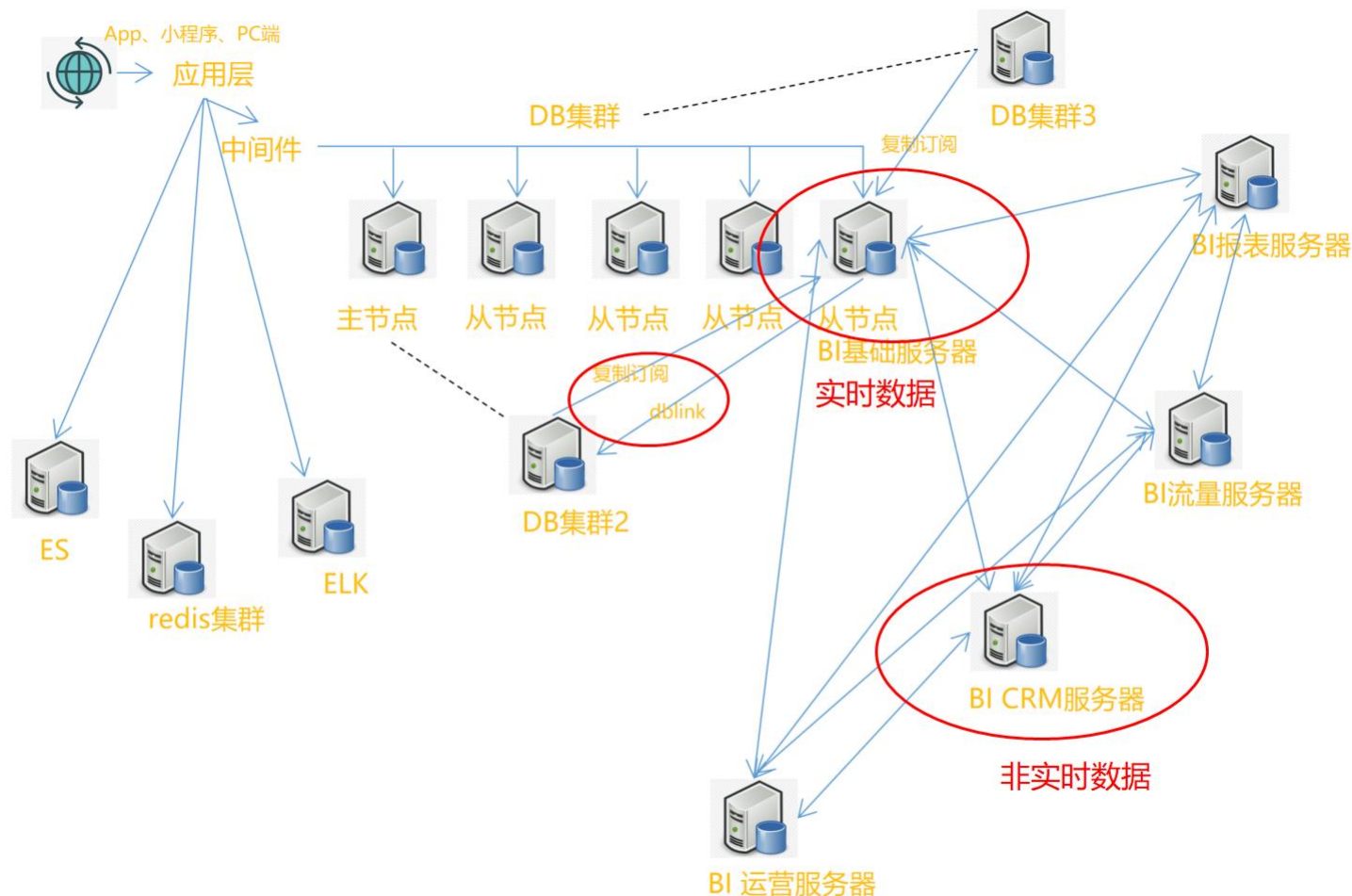
大量的复制订阅和dblink导致内耗很严重

数据量一大只能归档历史数据和分库分表

硬件性能有很大的瓶颈

通常一些大报表会跑几个小时

很难有扩展和提升空间



寻求解决方案

- 1、自建

考虑到自建的机器成本和人工成本，最终放弃

- 2、ADB

基于计算存储分离架构，支持海量数据实时写入可见及高性能在线分析。

支持按需弹性扩容计算资源和存储资源，提供数据冷热分层存储能力，以降低存储成本。

同时提供离线业务与在线业务的计算资源物理隔离的能力。

- 3、Hologres

一站式实时数仓，支持OLAP、Ad Hoc和高QPS点查。

支持海量数据实时写入、实时更新、实时分析，支持标准SQL。

云原生计算存储分离架构，多节点并行写入，并行计算。

支持高并发低延迟的在线数据服务（Serving），与MaxCompute、Flink、DataWorks深度融合，提供离在线一体化全栈数仓解决方案。

综合测试下来，Hologres性能略优于ADB，而且Hologres更加符合我们使用的业务场景，于是就选择了Hologres。

Part 3

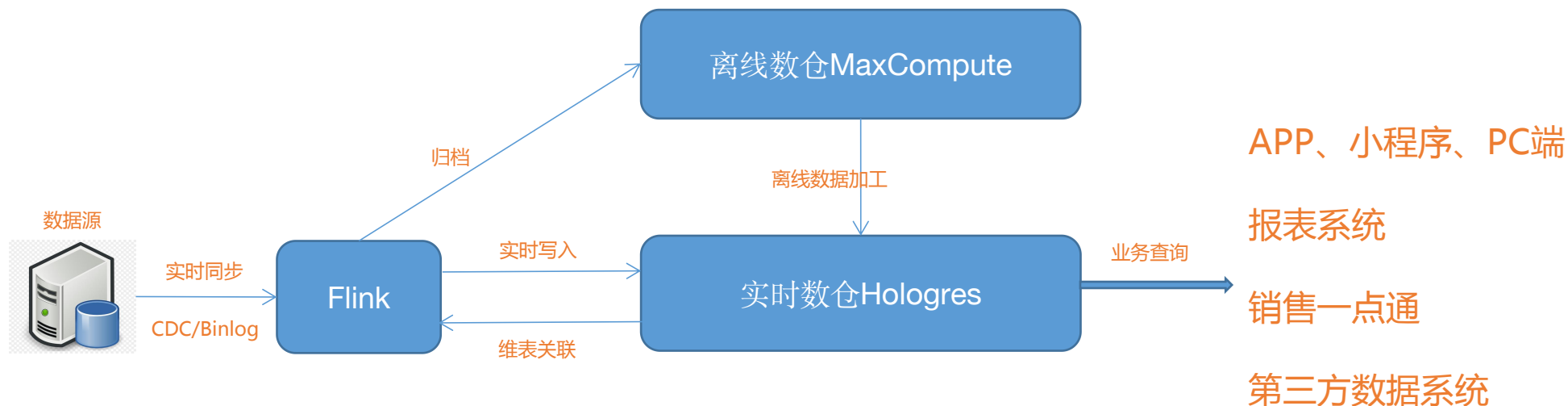
本来生活网大数据平台现状

本来生活网现在的大数据平台建设情况

亟待解决的痛点

- 1、提升查询效率，缩短出报表时间
- 2、可以同时处理更大量的数据
- 3、实时效率更高
- 4、更好更快的满足业务的各种各样的需求

从N到1、化繁为简、一站式实时数仓



一站式实时数仓理念： 分析服务一体化

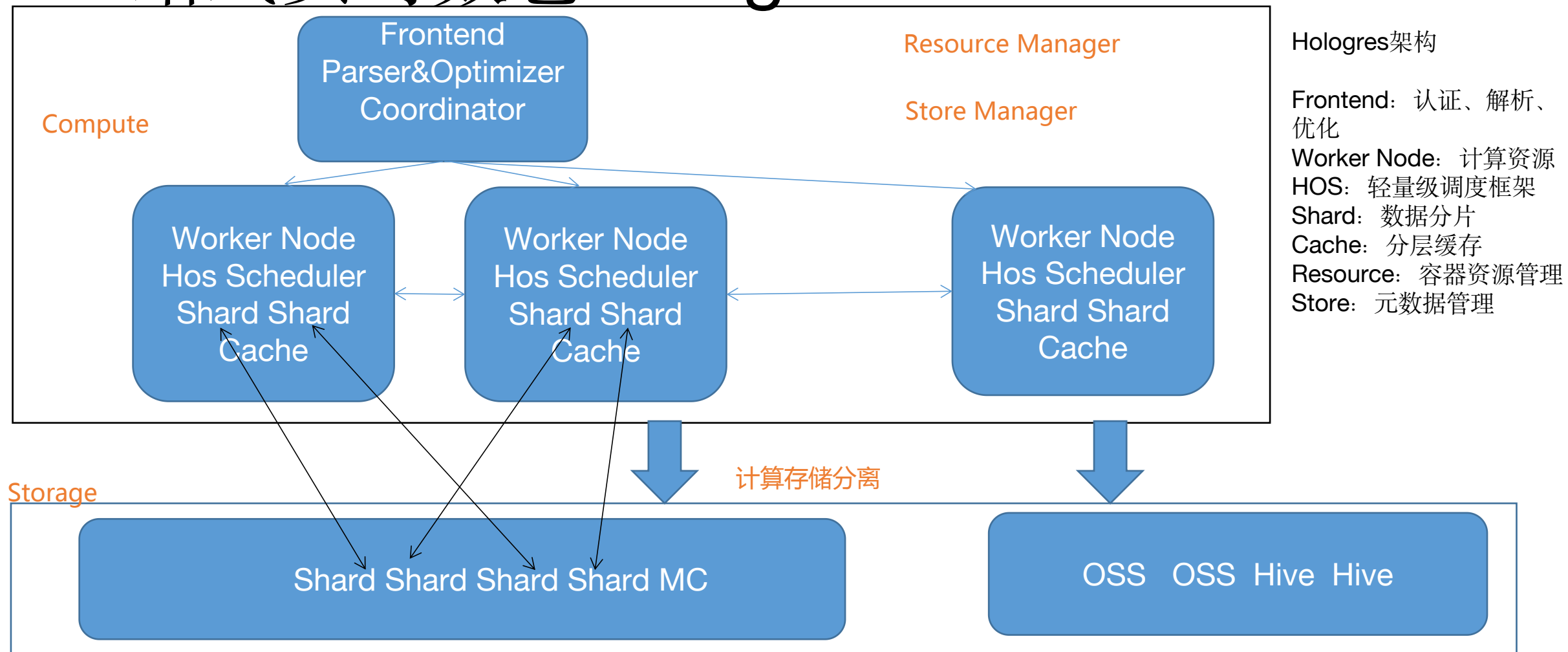
- 实时数仓Hologres:

统一数据存储： 离线数据批量导入

统一数据服务： 实时数据更新

数据应用： 多维分析、实时数据看板、在线业务

一站式实时数仓Hologres



解决痛点

- 1、查询速度加快
- 2、存储数据量更大
- 3、分析服务一体化
- 4、列存储对特定查询效率更高
- 5、数据分片可以并行处理更大的数据量

Part 4

本来生活网大数据平台展望

对本来生活网未来大数据平台建设的展望

未来展望

- 1、流批一体：
一套系统、一个逻辑
- 2、动态规则
智慧化营销

THANKS

TDDL

DistributedTable

DBproxy

HBase

PostgreSQL

SSD

MongoDB

Cassandra

GreatDB

Hyperbase

Hubble

DataCenter

VisualDataPlatform

Blockchain

ArgoDB

Distributed

DatabaseKernel

TemporalData

CloudnativeData

AIalgorithm