

行业研究/专题研究

2019年09月15日

行业评级:

计算机软硬件

增持(维持)

谢春生 执业证书编号: S0570519080006
研究员 xiechunsheng@htsc.com

郭雅丽 执业证书编号: S0570515060003
研究员 010-56793965
guoyali@htsc.com

郭梁良 021-28972067
联系人 guoliangliang@htsc.com

金兴
联系人 jinxing@htsc.com

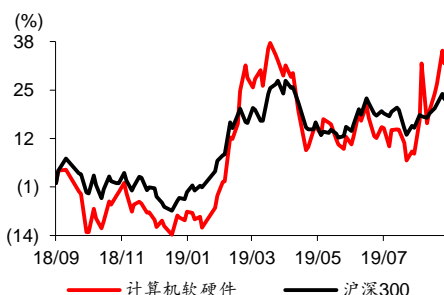
相关研究

1《计算机软硬件:网络治理规定利好舆情监控行业》2019.09

2《计算机软硬件:金融IT:内因和外因带来的投资机会》2019.09

3《计算机软硬件:趋势的落地和需求的分化》2019.09

一年内行业走势图



资料来源: Wind

我们对“Gauss 数据库”的理解

华为研究系列报告之一

数据库的重要性&华为推出新一代 Gauss 数据库

数据库是计算机行业的基础核心软件,所有应用程序的运行和数据处理都要与其进行数据交互。数据库的开发难度,不仅体现在与其他基础器件的适配,更在于如何实现对数据高效、稳定、持续的管理。Oracle、微软的数据库之所以能长久而不衰,一方面在于其强大的技术开发和产品升级迭代能力,另一方面在于其对数据库的 Knowhow 理解足够深,这个其他厂商短期难以超越的。华为在数据库领域逐步取得新的突破。2019 年华为推出了新一代的数据库产品 Gauss 数据库,该产品已经在金融、能源、政企等国内客户得到上线应用。

华为 Gauss 数据库: AI 原生&&支持异构计算

华为在数据库领域已经有 12 年的开发经验,从早期的摸索到现在的产品逐步成熟,中间也是经历了很多历程。华为的数据库产品系列命名为: GaussDB, 高斯数据库。华为 GaussDB 是一个企业级 AI-Native 分布式数据库。GaussDB 采用 MPP(Massive Parallel Processing)架构,支持行存储与列存储,提供 PB(Petabyte, 2 的 50 次方字节)级别数据量的处理能力。华为 Gauss 数据库是全球首款 AI-Native 数据库,能够同时支持 X86、ARM、GPU、NPU 等异构计算。

华为 Gauss 数据库: 三大产品线系列

GaussDB: 三大产品线系列。目前华为已经开发有三个产品系列: GaussDB 100、GaussDB 200、GaussDB 300。1) GaussDB 100: 主要以 OLTP 为主。目前该产品已经应用在招商银行。2) GaussDB 200: 以 OLAP 为主,兼顾 OATP。该产品目前正在工商银行得到上线应用。3) GaussDB 300: HTAP, 是企业级分布式 HTAP 数据库 (Hybrid Transaction and Analytical Process, 混合事务和分析处理)。

华为 Gauss 数据库: 华为 IT 生态体系不可或缺

从生态体系来看, Oracle 数据库成为全球第一数据库的地位,也是经过了很多次版本的更新升级,更与上世纪 80 年代开始的全球 IT 生态体系的逐步确立有关。Oracle 数据库世界霸主地位,是随着 Windows 操作系统、Intel X86 芯片一起建立的 PC 时代的 IT 底层生态而逐步确立的。而 AWS 数据库则是适应了互联网时代(云计算时代)新的计算场景对数据库的新需求,再加上自身的云生态体系,逐步迎来了客户使用的推广。从华为 Gauss 数据库来看,华为 IT 架构的底层生态已经逐步建立起来,包括芯片、操作系统、数据库等,这些在华为 IT 体系内部是高度耦合的。

风险提示: 数据库技术升级速度低于预期的风险, 数据库产品应用推广低于预期的风险, 数据库应用场景拓展低于预期风险。

正文目录

华为数据库分析.....4

 华为 DB 开发历程.....4

初识华为 GaussDB6

详解华为 GaussDB 11

 华为 GaussDB 200 11

 华为 GaussDB 300 18

数据库产品对比分析21

 Oracle 数据库.....21

 AWS 数据库25

 华为 GaussDB：对比.....28

总结.....31

 风险提示.....31

图表目录

图表 1: 数据库的演进.....	4
图表 2: 华为 GaussDB 对 AI 技术的应用.....	5
图表 3: 华为 GaussDB 特点	6
图表 4: 华为 GaussDB 产品线.....	7
图表 5: OLTP&OLAP 对比.....	8
图表 6: 华为 GaussDB OLTP&OLAP 特点	9
图表 7: 华为 GaussDB HTAP 特点	10
图表 8: 数据库: 行式存储	11
图表 9: 数据库: 列式存储	12
图表 10: 数据库: 行式&列式存储对比	12
图表 11: 华为 GaussDB 200 架构.....	14
图表 12: 华为 GaussDB 200 软件构成	16
图表 13: 华为 GaussDB 特点	16
图表 14: 华为 GaussDB 200 架构	17
图表 15: 华为 GaussDB200 数据查询过程.....	18
图表 16: 华为 GaussDB 300 集群架构	20
图表 17: Oracle 数据库不断演进	21
图表 18: Oracle 数据库 18C 的升级	22
图表 19: Oracle 数据库 19C 的升级	22
图表 20: Oracle 数据库支持多种数据类型.....	23
图表 21: Oracle 的分布式数据库	24
图表 22: Oracle 数据库支持多种数据类型.....	25
图表 23: Amazon Aurora 典型客户	27
图表 24: Amazon DynamoDB 典型客户	28
图表 25: 华为 GaussDB 200 典型客户	29

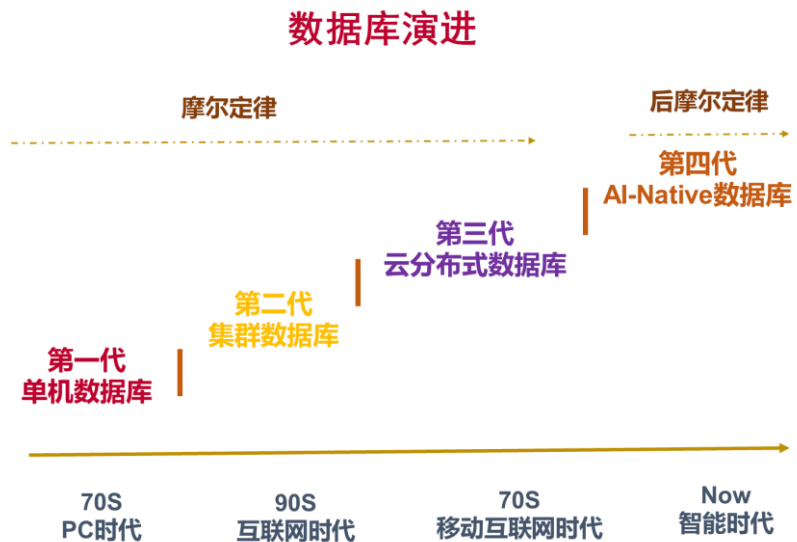
华为数据库分析

数据库是计算机行业的基础核心软件，所有应用程序的运行和数据处理都要与其进行数据交互。2008 年阿里提出“去 IOE”，而 10 年之后，我们现在来看，发现 Oracle 的数据库是最难替换的。不仅是因为 Oracle 的数据库沉淀了大量的企业客户数据，更是因为数据库产品开发难度确实比较大。数据库的开发难度不亚于操作系统，属于整个 IT 架构的基础软件（数据库软件在操作系统之上，我们可以将其称为类中间层的基础软件）。而且数据库的开发需要与底层计算架构高度相关和耦合，是适配 X86 架构，还是适配 ARM 架构等等。

当然以上这些都是数据库的与其他基础器件的适配，数据库难度更大的地方在于如何实现数据高效、稳定、持续的管理。Oracle、微软的数据库之所以能长久而不衰，一方面在于其强大的技术开发和产品升级迭代能力，另一方面在于其对数据库的 Knowhow 理解足够深，这个其他厂商短期难以超越的。

回到这篇文章的主题：华为数据库。华为在 IT 的底层架构，逐步搭建起自己的基础架构，建立华为生态。我们这次把华为数据库进行讲解，并对目前主流的数据库进行对比。只有对比，才能发现不同。

图表1：数据库的演进



资料来源：华为官网，华泰证券研究所

华为 DB 开发历程

华为对数据库的开发经历了长达 12 年左右的时间。

2007 年，华为开始着手研发内存数据库，项目代号为 GMDB。这个项目的背景是，当时电信实施实时计费，电信行业对数据库有特殊的要求，有些需要定制化开发。而当时国外的数据库产品主要是标准化产品。为了满足客户需求，华为当时开始研发内存数据库。

2010 年，华为开始从内存数据库向通用关系型数据库进行拓展，逐步将非内存数据库的功能融入到数据库产品中。

2012 年，华为数据库性能得到显著提升，GMDB 开始逐步商用化，主要应用于电信计费。同时，该产品也在华为内部的部分部门开始使用。

2013 年，华为 OLTP 数据库开始上线（后面我们会详细介绍 OLTP 和 OLAP）。

2014 年，华为开发出第一个 OLAP 数据库版本（OLAP 我们可以简单理解为：是针对大量数据的分析型数据库）。

2015 年，华为与工商银行一起联合研发。Gauss OLAP 数据库在工商银行上线，逐步替代海外的数据仓库。

2017 年，华为与招商银行一起联合开发 GaussDB。同时，华为启动面向事务和分析混合处理的数据库开发，即 HTAP。

2018 年，华为 Gauss OLTP 数据库（事务型数据库）开始在招商银行综合支付交易系统成功上线。承接招商银行“手机银行”和“掌上生活”两大 App 交易流水流量。

2018 年，Gauss HTAP 数据库推出，并在民生银行得到应用。

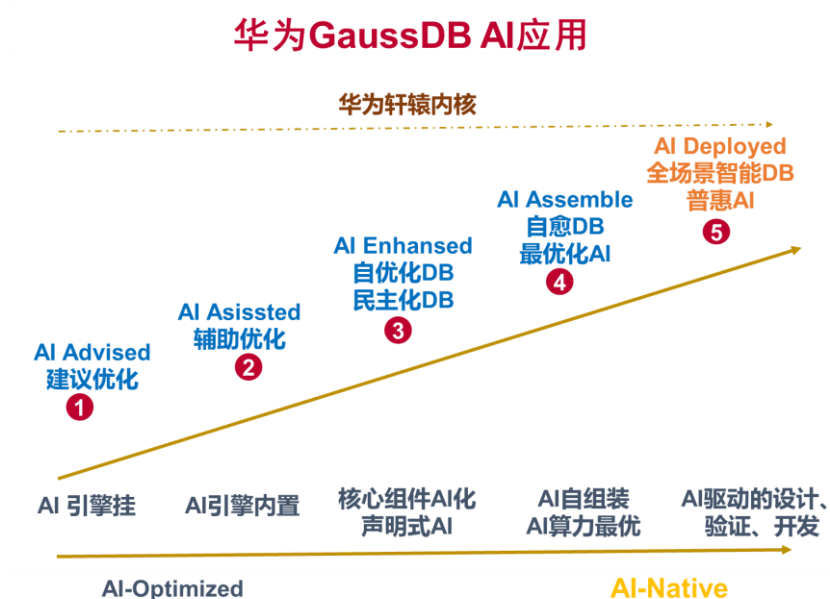
从华为 Gauss 数据库产品演化至今来看，

1) 华为数据库产品的研发是从内存数据库开始，逐步向通用关系型数据库延伸，这与 Oracle、AWS 数据库开发的起点并不完全一样。

2) 华为数据库产品类型，包括了 OLTP、OLAP，同时还研发出 HTAP 产品。我们认为，从产品应用角度来看，华为 OLAP（分析型数据库）大规模应用的时点更早一些。Oracle 的 OLTP（事务型数据库）在全球领域的竞争优势非常明显，这一领域的数据库产品比较难替代。

3) 华为的 OLTP 数据库是通过与大客户合作，特别是银行大客户合作（工商银行、招商银行），来不断进行产品迭代和完善的。我们认为，这也是华为数据库能够快速成长的主要原因。

图表2： 华为 GaussDB 对 AI 技术的应用



资料来源：华为官网，华泰证券研究所

初识华为 GaussDB

华为在数据库领域已经有 12 年的开发经验，从早期的摸索到现在的产品逐步成熟，中间也是经历了很多历程。目前，华为数据库逐步建立起三大产品系列。

华为的数据库产品系列命名为：GaussDB，高斯数据库。高斯，是德国伟大的数学家，近代数学的奠基者之一，高斯、阿基米德、欧拉、牛顿被世人称为世界上最伟大的四位数学家。华为将自己的数据库命名为 Gauss 系列，也有向数学致敬的意味。

GaussDB：开源数据库。华为的 Gauss 数据库是一个开源数据库，基于 PostgreSQL9.2 开发。我们知道 PostgreSQL 本身就是一个开源数据库品牌。现在除了 Oracle DB、微软的 SQL Server 等传统老牌数据产品之外，目前新开发的数据库产品，开源数据库占比较大的部分。包括我们看到的 AWS 的 Aurora 数据库、阿里的飞天数据库、华为的 Gauss 数据库，以及数据库新进入者 MongoDB 等。

GaussDB：分布式&AI 原生。华为 GaussDB 是一个企业级 AI-Native 分布式数据库。GaussDB 采用 MPP(Massive Parallel Processing)架构，支持行存储与列存储，提供 PB(Petabyte, 2 的 50 次方字节)级别数据量的处理能力。可以为超大规模数据管理提供高性价比的通用计算平台，也可用于支撑各类数据仓库系统、BI(Business Intelligence)系统和决策支持系统，为上层应用的决策分析提供服务。华为 GaussDB 将 AI 能力植入到数据库内核的架构和算法中，为用户提供更高性能、更高可用、更多算力支持的分布式数据库。

图表3： 华为 GaussDB 特点



资料来源：华为官网，华泰证券研究所

GaussDB：三大产品线系列。高斯数据库研发始于 2011 年。目前已经开发有三个产品系列：GaussDB 100、GaussDB 200、GaussDB 300。

GaussDB 100：主要以 OLTP 为主。GaussDB 100 研发开始于 2011 年，与后面的 GaussDB 200/300 不同，GaussDB 100 并不是一个分布式数据库。GaussDB 100 包括两条线，一条产品线是基于单机版开源数据库 PostgreSQL 研发的产品，另一条线是自研内核的 GaussDB 100 产品。后面这一条线是近几年华为研发的产品。目前该产品已经应用在招商银行。

GaussDB 100 主要是 OLTP，即事务型数据库。

GaussDB 200: 以 OLAP 为主，兼顾 OLTP。华为 GaussDB 200 开始于 2012 年，在基于传统关系型数据库的 SQL 引擎和事务强一致性等基础上，进行了分布式、并行计算的改造。历时 6 年，打造了一款架构领先的分析型数据库，为各行业 PB 级海量数据分析提供有竞争力的解决方案。GaussDB 200 既可以适用于 OLTP，也可以应用于 OLAP。

GaussDB 300: HTAP, OLTP 和 OLAP。GaussDB 300 是一个分布式并行关系型数据库系统，是企业级分布式 HTAP 数据库（Hybrid Transaction and Analytical Process，混合事务和分析处理）。GaussDB 300 架构上着重构筑传统数据库的企业级能力和互联网分布式数据库的高扩展和高可用能力，完全兼容 SQL 标准，提供百万级 TPMC 的交易处理能力和企业级可靠性。

GaussDB 200/300 都是基于开源数据库 PostgreSQL 研发，虽然是基于开源数据库，但已经对开源代码进行了大量修改，在很大程度上接近于自研。GaussDB 200/300 既可以支持 OLTP 也可以支持 OLAP，也是华为投入精力最大、研发时间最长的产品线。目前已经在工商银行和民生银行应用。

图表4： 华为 GaussDB 产品线



资料来源：华为官网，华泰证券研究所

在以上我们对华为 GaussDB 的介绍当中，提到了数据库领域比较重要的两个概念：OLTP 和 OLAP。下面我们就介绍下这两个概念，以及其所对应的数据库类型。

华为 GaussDB 数据库包括：事务性（OLTP）数据库、分析型（OLAP）数据库和混合负载（HTAP）数据库。这里需要解释下 OLTP、OLAP、HTAP 之间的区别，这也是数据库最基本的内容。

数据库系统一般分为两种类型：一种是面向前台应用的，应用比较简单，但是重吞吐和高并发的 OLTP 类型；一种是重计算的，对大数据集进行统计分析的 OLAP 类型。

1) OLTP: 联机事务处理 OLTP (on-line transaction processing)

它是事件驱动、面向应用的，比如电子商务网站的交易系统就是典型的 OLTP 系统。OLTP 的基本特点是：

数据在系统中产生；
基于交易的处理系统 (Transaction-Based)；
每次交易牵涉的数据量很小；
对响应时间要求非常高；
用户数量非常庞大，主要是操作人员；
数据库的各种操作主要基于索引进行。

2) OLAP: 联机分析处理 OLAP (On-Line Analytical Processing)

是基于数据仓库的信息分析处理过程，是数据仓库的用户接口部分。OLAP 系统是跨部门的、面向主题的，其基本特点是：

本身不产生数据，其基础数据来源于生产系统中的操作数据 (Operational Data)；
基于查询的分析系统；
复杂查询经常使用多表联结、全表扫描等，牵涉的数量往往十分庞大；
响应时间与具体查询有很大关系；
用户数量相对较小，其用户主要是业务人员与管理人员；
由于业务问题不固定，数据库的各种操作不能完全基于索引进行。

OLTP 是传统的关系型数据库的主要应用，主要是基本的、日常的事务处理，例如银行交易。OLAP 是数据仓库系统的主要应用，支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。

OLTP 系统强调数据库内存效率，强调内存各种指标的命令率，强调绑定变量，强调并发操作。OLAP 系统则强调数据分析，强调 SQL 执行市场，强调磁盘 I/O，强调分区等。

图表5: OLTP&OLAP 对比

	OLTP	OLAP
用户	操作人员，低层管理人员	决策人员，高级管理人员
功能	日程操作处理	分析决策
DB 设计	面向应用	面向主题
数据	当前的，最新的，细节的，二维的，分离的	历史的，聚焦的，多维的，集成的，统一的
存取	读/写数十条记录	读上百万条记录
工作单位	简单的事务	复杂的查询
用户数	上千个	上百万个
DB 大小	100MB-GB	100GB-TB
时间要求	具有实时性	对时间的要求不严格
主要应用	数据库	数据仓库

资料来源：至顶网，华泰证券研究所

3) HTAP: 混合事务和分析处理 (Hybrid Transaction and Analytical Process)，既可以应用于事务型数据库场景，亦可以应用于分析型数据库场景。

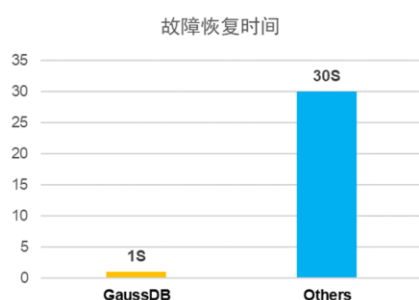
GaussDB OLTP 数据库，业界首创 Switch Turbo 技术，AZ 内 TRO<1 秒，满足金融场景下的数据库高可用诉求。

GaussDB OLAP 数据库，可以帮助客户实现 PB 级海量数据高效分析，目前已经广泛应用于金融、运营商、政府等行业。

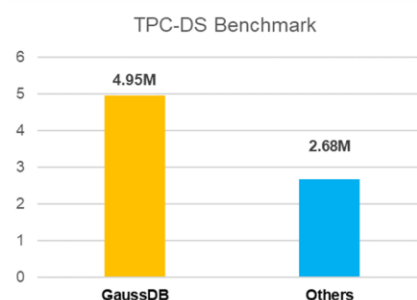
图表6： 华为 GaussDB OLTP&OLAP 特点

华为 GaussDB 特性

OLTP数据库
故障恢复业内最快



OLAP数据库
TPC-DS测试性能业界最高



资料来源：华为官网，华泰证券研究所

GaussDB HTAP 数据库，多模引擎支持五种数据类型融合处理，包括流、图、空间、文本、结构化，可以解决集中式架构扩展性和性能瓶颈问题，同时分散风险，提升业务连续性。

图表7： 华为 GaussDB HTAP 特点

华为 GaussDB HTAP 特性



资料来源：华为官网，华泰证券研究所

华为 GaussDB 值得关注的点：

1、全球首款 AI-Native 数据库。

AI 原生数据库是 GaussDB 的主要特点之一。华为将 AI 引擎内置到 GaussDB 全系产品中，使其具备一定的自运维、自管理、自调优、故障自诊断和自愈的能力。华为也希望把在芯片、算法上面的优势，集中体现到数据库上来。

客观来讲，其实对于在数据库中植入 AI 技术，并不是一个新鲜做法。Oracle 在几个版本之前就开始就植入了 AI 技术，开启了“Autonomous”之旅。

2、异构计算支持 X86、ARM、GPU、NPU。

这个也是 Gauss 数据库与其他数据库比较大的不同。目前主流的数据库产品，包括 OracleDB、MySQL、SQL Server 等，基本都是支持 X86 架构。我们认为，华为数据库对于异构计算的支持，可能是为该数据库未来向更多计算场景的应用做准备。我们知道，5G 带来计算场景的变革或将更大。

详解华为 GaussDB

华为的 Gauss 数据库现在推广的产品主要是 GaussDB 100、GaussDB 200 和 GaussDB 300。我们这里主要对 GaussDB 200 和 GaussDB 300 这两个系列产品进行介绍和解读。

华为 GaussDB 200

1、GaussDB 200 简介

GaussDB 200 是企业级的大规模并行处理关系型数据库，采用 MPP（Massively Parallel Processing）架构，支持行存储与列存储，提供 PB（Petabyte，2 的 50 次方字节）级别数据量的处理能力。

从以上对 GaussDB 的描述中，我们至少能够理解到以下几层意思：

- 1) GaussDB 200 是一个关系型数据库，不是 No-SQL 数据库。
- 2) 它利用了分布式并行处理技术。早期传统数据库并不是分布式架构。分布式并行架构更适合于处理互联网高并发数据。
- 3) 支持行存储和列存储。

这里需要解释下，数据库领域中的行存储和列存储区别。

行存储（Row-based）：对于传统的关系型数据库，比如甲骨文的 OracleDB 和 MySQL，IBM 的 DB2、微软的 SQL Server 等，一般都是采用**行存储（Row-based）**行。在基于行式存储的数据库中，数据是按照行数据为基础逻辑存储单元进行存储的，一行中的数据在存储介质中以连续存储形式存在。

图表8： 数据库：行式存储

数据库：行式存储

传统行式数据库（关系型数据库）

	1	2	3	4	5	6	7	8	9	...
R1
R2
R3
R4
R5

- 1、数据是按行存储的
- 2、没有索引的查询使用大量I/O
- 3、建立索引和物化视图需要花费大量时间和资源
- 4、面对查询的需求，数据库必须被大量膨胀才能满足性能需求

资料来源：至顶网，华泰证券研究所

列式存储 (Column-based) 是相对于行式存储来说的, 新兴的 Hbase、HP Vertica、EMC Greenplum 等分布式数据库均采用列式存储。在基于列式存储的数据库中, 数据是按照列为基础逻辑存储单元进行存储的, 一列中的数据在存储介质中以连续存储形式存在。

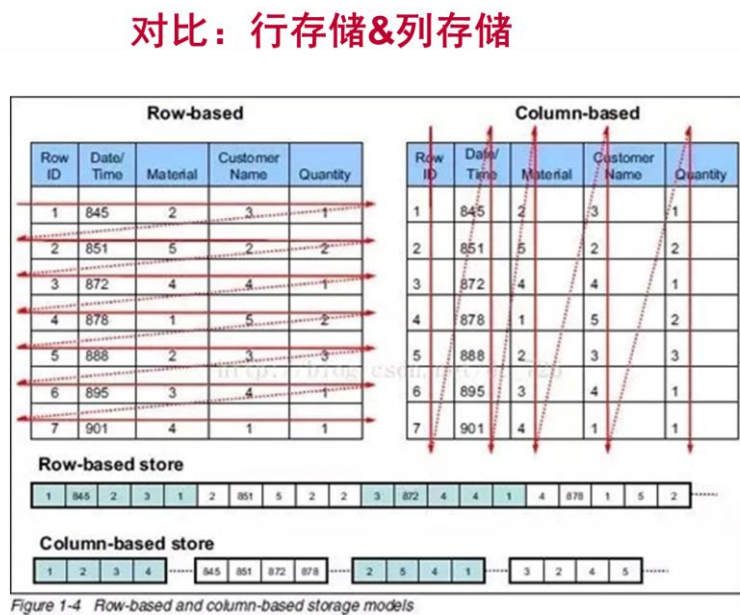
图表9: 数据库: 列式存储



资料来源: 至顶网, 华泰证券研究所

传统的行式数据库, 是按照行存储的, 维护大量的索引和物化视图无论是在时间 (处理) 还是空间 (存储) 面成本都很高。而列式数据库恰恰相反, 列式数据库的数据是按照列存储, 每一列单独存放, 数据即是索引。只访问查询涉及的列, 大大降低了系统 I/O, 每一列由一个线索来处理, 而且由于数据类型一致, 数据特征相似, 极大方便压缩。

图表10: 数据库: 行式&列式存储对比



资料来源: 至顶网, 华泰证券研究所

第一, 从存储角度来看, 对于列式存储来说, 一行数据包含一个列或者多个列, 每列有单独一个 cell 来存储数据。而行式存储, 则是把一行数据作为一个整体来存储。另外, 列式存储天生就是适合压缩, 因为同一列里面的数据类型基本是相同。

第二，从查询角度来看，行式存储比较适合随机查询，而且，关系型数据库（RDBMS）大多提供二级索引，在整行数据的读取上，要优于列式存储。但是，行式存储不适合扫描，这意味着要查询一个范围的数据，行式存储需要扫描整个表

基于以上，我们可以看出，GaussDB 200 兼具了行存储和列存储的优势。

4) MPP 架构。这个与上面的分布式并行是一体的。在数据库中，MPP 架构并不是新技术。目前 Oracle 的数据库，包括国内的南大通用的分析型数据库（OLAP）都采用了 MPP 架构。

MPP (Massively Parallel Processing)，即大规模并行处理，在数据库非共享集群中，每个节点都有独立的磁盘存储系统和内存系统，业务数据根据数据库模型和应用特点划分到各个节点上，每台数据节点通过专用网络或者商业通用网络互相连接，彼此协同计算，作为整体提供数据库服务。非共享数据库集群有完全的可伸缩性、高可用、高性能、优秀的性价比、资源共享等优势。

简单来说，MPP 是将任务并行的分散到多个服务器和节点上，在每个节点上计算完成后，将各自部分的结果汇总在一起得到最终的结果。

说起 MPP，我们不得不说大数据处理中应用到的另一个技术：Hadoop。这个词曾经在 2016 年中国大数据行业比较流行。我们来看下，MPPDB、Hadoop 与传统数据库技术在应用场景上的对比。

MPPDB 与 Hadoop 都是将运算分布到节点中独立运算后进行结果合并（分布式计算），但由于依据的理论和采用的技术路线不同而有各自的优缺点和适用范围。两种技术以及传统数据库技术的对比如下：

Hadoop 和 MPP 两种技术的特定和适用场景为：

Hadoop 在处理非结构化和半结构化数据上具备优势，尤其适合海量数据批处理等应用要求。

MPP 适合替代现有关系数据库机构下的大数据处理，具有较高的效率。

MPP 适合多维度数据自助分析、数据集市等；Hadoop 适合海量数据存储查询、批量数据 ETL、非机构化数据分析(日志分析、文本分析)等。

由上述对比可预见未来大数据存储与处理趋势：MPPDB+Hadoop 混搭使用，用 MPP 处理 PB 级别的、高质量的结构化数据，同时为应用提供丰富的 SQL 和事务支持能力；用 Hadoop 实现半结构化、非结构化数据处理。这样可以同时满足结构化、半结构化和非结构化数据的高效处理需求。

GaussDB 200 在核心技术上跟传统数据库相比有巨大优势，可以解决很多行业用户的数据处理性能问题，可以为超大规模数据管理提供高性价比的通用计算平台，并可用于支撑各类数据仓库系统、BI（Business Intelligence）系统和决策支持系统，统一为上层应用的决策分析等服务。

2、GaussDB 200 应用场景

GaussDB 200 面向行业大数据应用，可以适用于以下场景：

1) 详单查询

GaussDB 200 具备 PB 级数据负载能力，通过内存分析技术满足海量数据边入库边查询要求，适用于安全、电信、金融、物联网等行业的详单查询业务。

从这点可以看出，这一功能的实现是发挥了 GaussDB 200 在 OLTP 方面能力。

2) 数据仓库

GaussDB 200 具备百 TB 级数据支撑能力，可以高效处理百亿行多表连接查询，适用于操作数据存储 ODS（Operational Data Store）、企业数据仓库 EDW（Enterprise Data Warehouse）、数据集市 DM（Data Mart）。

以上这一功能体现了 GaussDB 200 在 OLAP 方面的能力。

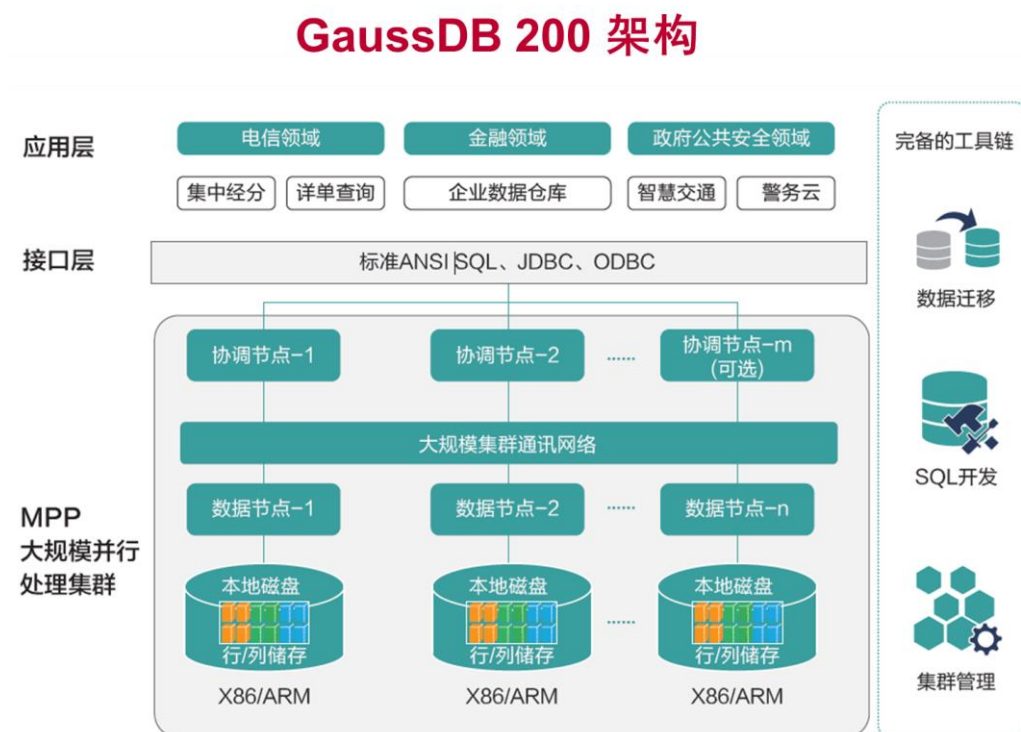
3) 混合负载

GaussDB 200 基于海量数据查询统计分析能力与事务处理能力，行列混存技术同时满足联机事务处理 OLTP（On-Line Transaction Processing）与联机分析处理 OLAP（Online Analytical Processing）混合负载场景。

4) 大数据分析

支持结构化数据 PB 级分析能力。分布式并行数据库集群满足 PB 级结构化大数据的分析能力。

图表11： 华为 GaussDB 200 架构



资料来源：华为官网，华泰证券研究所

GaussDB 200 是一个分布式并行关系型数据库系统。提供了以下功能：

1) 标准 SQL 支持

支持标准的 SQL-92/SQL:1999/SQL:2003 规范，支持 GBK 和 UTF-8 字符集，支持 SQL 标准函数与分析函数，支持存储过程。

2) 数据库存储管理功能

支持表空间，支持在线扩容功能。

3) 提供组件管理和数据节点 HA (High Availability)

支持数据库事务 ACID 特性（即原子性 Atomicity、一致性 Consistency、隔离性 Isolation 和持久性 Durability），支持单节点故障恢复，支持负载均衡等。

4) 应用程序接口

支持标准 JDBC 4.0 的特性和 ODBC 3.5 特性。

5) 安全管理

支持 SSL 安全网络连接、用户权限管理、密码管理、安全审计等功能，保证数据库在管理层、应用层、系统层和网络层的安全性。

3、GaussDB 200 技术特点

GaussDB 200 具有低成本、高性能、高可靠和支持海量数据存储的特点。

1) 低成本

基于分布式 x86 架构，客户硬件投资成本低。

支持标准的 SQL92、SQL99、SQL2003 规范，支持客户应用系统平滑迁移。

2) 高性能

行列混合存储引擎，数据按照最优负载模型选择存储方式，性能更优。

支持基于服务等级协议 SLA (Service-Level Agreement) 策略的负载管理，保障并发任务的服务质量。

支持基于代价的查询优化模型，提升复杂查询性能。

分布式、并行化的查询处理模型，充分利用系统计算资源和 IO 资源。

支持并行数据导出和导入。

3) 高可靠

硬件级高可靠：磁盘 Raid、交换机堆叠及网卡 bond、不间断电源 UPS (Uninterruptible Power Supply)。

件级高可靠：集群 CM、CN、GTM、DN 实例全方位 HA。

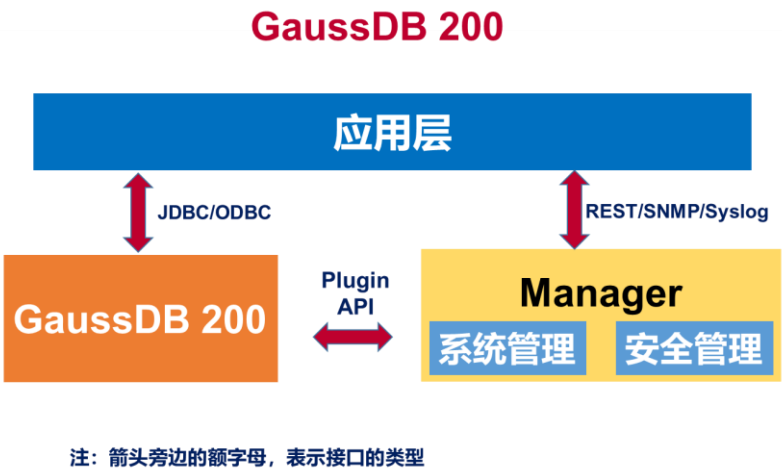
数据存储安全可靠：在安全认证的基础上，支持数据在数据库内的加密存储，防止三方人员绕过数据库认证机制直接读取数据文件中的数据。

4) 支持海量数据

集群最大可扩展至 2048 个节点，支撑 PB 级数据管理能力。

集群规模按用户需求弹性伸缩，扩展业务不中断，减少用户投资成本。

图表12： 华为 GaussDB 200 软件构成



资料来源：华为官网，华泰证券研究所

各组件提供功能如下：

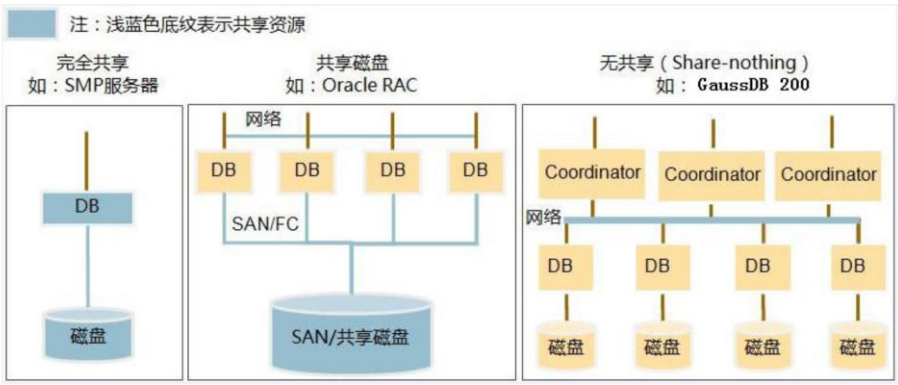
GaussDB 200：基于 MPP 架构的新型数据库，为 GaussDB 200 提供 PB 级数据负载能力、百 TB 级数据支撑能力、海量数据查询统计分析能力与事务处理能力、支持结构化数据 PB 级分析能力等。

Manager：作为运维系统，负责 GaussDB 200 的集群管理，支持大规模集群的安装部署、监控、告警、用户管理、权限管理、审计、服务管理、健康检查、问题定位、升级和补丁等。

GaussDB 200 采用 Share-nothing 架构，由多个拥有独立且互不共享 CPU、内存、存储等系统资源的节点组成。在这样的系统架构中，业务数据被分散存储在多个物理节点上，数据分析任务被推送到数据所在位置就近执行，通过控制模块的协调，并行地完成大规模的数据处理工作，实现对数据处理的快速响应。Share-nothing 又称为无共享架构。

图表13： 华为 GaussDB 特点

GaussDB 架构对比



资料来源：华为官网，华泰证券研究所

Share-nothing 架构具备如下优点：

1) 最易于扩展的架构

为商业智能 BI (Business Intelligence) 和数据分析的高并发、大数据量计算提供按需扩展的能力；

自动化的并行处理机制。

2) 内部自动并行处理，无需人工分区或优化

数据加载与访问方式与一般数据库相同；

数据分布在所有的并行节点上；

每个节点只处理其中一部分数据。

3) 最优化的 I/O 处理

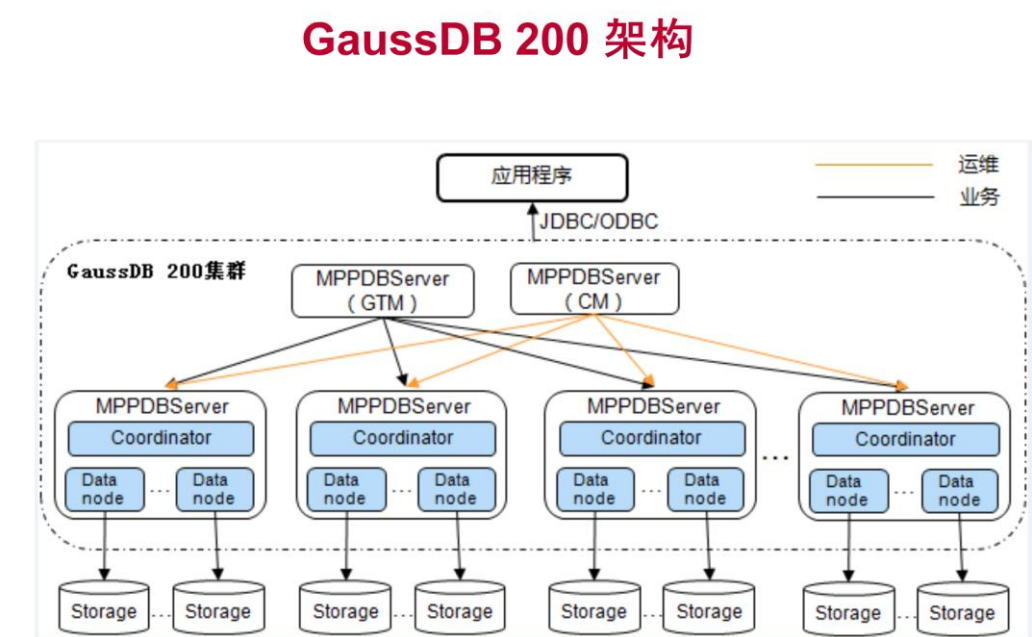
所有的节点同时进行并行处理；

节点之间完全无共享，无 I/O 冲突。

4) 增加节点实现存储、查询及加载性能的线性扩展

GaussDB 200 由多个 MPPDBServer 组成，GaussDB 200 结构具体如图所示。

图表14： 华为 GaussDB 200 架构



资料来源：华为官网，华泰证券研究所

作为关系型数据库系统，GaussDB 200 主要业务为数据的查询与存储。GaussDB 200 进行数据查询的流程如图 15 所示。

具体查询流程如下：

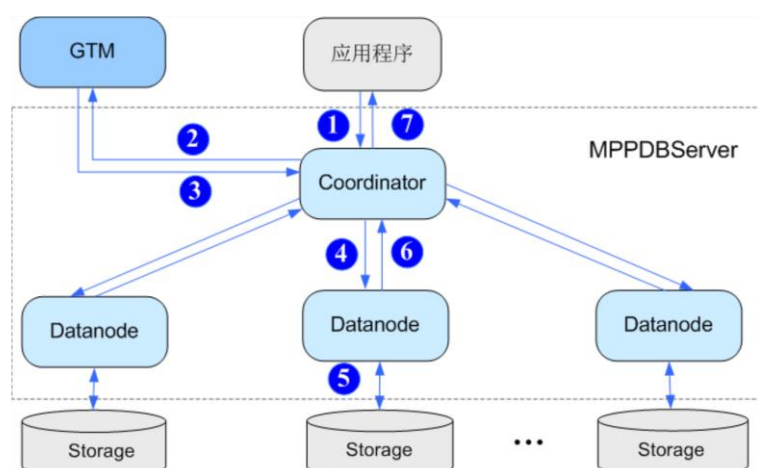
- 1) 用户通过应用程序发出查询本地数据的 SQL 请求到 Coordinator。
- 2) Coordinator 接收用户的 SQL 请求，分配服务进程，向 GTM 请求分配全局事务信息。
- 3) GTM 接收到 Coordinator 的请求，返回全局事务信息给 Coordinator。
- 4) Coordinator 根据数据分布信息以及系统元信息，解析 SQL 为查询计划树，从查询计划树中提取可以发送到 Datanode 的执行步骤，封装成 SQL 语句或者子执行计划树，发送到 Datanode 执行。

- 5) Datanode 接收到读取任务后，查询具体 Storage 上的本地数据块。
- 6) Datanode 任务执行后，将执行结果返回给 Coordinator。
- 7) Coordinator 将查询结果通过应用程序返回给用户。

GaussDB 200 数据存储流程与数据查询流程相近，请参考数据查询流程，此处不再介绍。

图表15： 华为 GaussDB200 数据查询过程

GaussDB 200 数据查询流程



资料来源：华为官网，华泰证券研究所

华为 GaussDB 300

GaussDB 300 是企业级分布式 HTAP 数据库 (Hybrid Transaction and Analytical Process, 混合事务和分析处理)，架构上着重构筑传统数据库的企业级能力和互联网分布式数据库的高扩展和高可用能力，完全兼容 SQL 标准，提供百万级 TPMC 的交易处理能力和企业级可靠性。

GaussDB 300 是一个分布式并行关系型数据库系统。提供了以下功能：

1) 标准 SQL 支持

支持标准的 SQL92/SQL99/SQL2003 规范，支持 GBK 和 UTF-8 字符集，支持 SQL 标准函数与分析函数，支持存储过程。

2) 数据库存储管理功能

支持表空间，支持在线扩容功能。

3) 提供组件管理和数据节点 HA (High Availability)

支持数据库事务 ACID 特性 (即原子性 Atomicity、一致性 Consistency、隔离性 Isolation 和持久性 Durability)，支持单节点故障恢复，支持负载均衡等。

4) 应用程序接口

支持标准 JDBC 4.0 的特性和 ODBC 3.5 特性。

5) 安全管理

支持 SSL 安全网络连接、用户权限管理、密码管理、安全审计等功能，保证数据库在管理层、应用层、系统层和网络层的安全性。

GaussDB 300 具备如下特性：

1) 高扩展：容量和性能按需水平扩展

无共享架构，性能线性扩展能力。

应用透明的 Sharding 能力。

2) 高可用：双活和两地三中心高可用

读写分离和两地三中心部署。

分片双活和两地三中心部署。

3) 高性能：支持高吞吐强一致性事务能力

高并发 MVCC 和日志管理。

跨节点高性能、强一致事务。

4) 易管理：易迁移，易监控，运维

兼容 SQL2003，存储过程等高级功能。

在线升级，在线扩容。

GaussDB 300 的分布式能力做到对业务透明，能替换传统 OLTP 数据库应用，也能替换数据库中间件、数据库分库分表等 Sharding 方案。同时它也让开发运维人员不用关注数据库数据分布和扩展的细节问题，专注于业务开发，极大的提升研发的生产力。

GaussDB 300 应用场景

GaussDB 300 面向行业大数据应用，可以适用于以下场景：

1) 核心交易处理

具备大并发、百万级 TPMC 事务处理能力，提供集群高可用、同城双活和跨城市级容灾能力，能够支撑银行核心账务、渠道交易、互联网金融等关键 OLTP 应用场景。

2) 复杂事务处理

支持标准 SQL92、SQL99 和 SQL2003，提供存储过程、触发器、分页、游标等能力，满足 CRM、ERP 等典型企业应用场景的复杂事务能力要求。

3) 混合负载场景

基于行列混存、动态负载管理和多租户等技术，具备复杂事务处理和海量数据查询统计分析能力，同时满足联机事务处理 OLTP (On-Line Transaction Processing) 与联机分析处理 OLAP (Online Analytical Processing) 混合负载场景。

GaussDB 300 技术特点

华为 GaussDB 300 具有如下技术特点：

1) 低成本

基于分布式 x86 架构（支持鲲鹏架构），客户硬件投资成本低。

支持标准的 SQL92/SQL99/SQL2003 规范，支持客户应用系统平滑迁移。

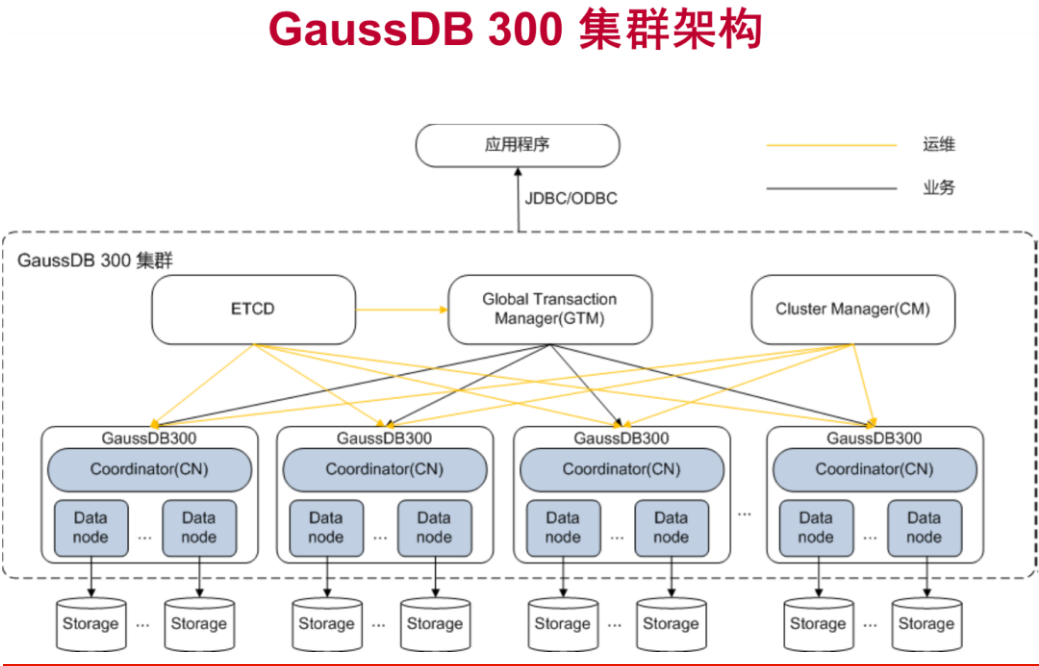
2) 高性能

通过列存储、向量化执行引擎、分布式查询处理等关键技术，实现百亿数据量查询秒级响应。

3) 支持海量数据

集群最大可扩展至 32 个节点，支撑 TB 级数据管理能力。
集群规模按用户需求弹性伸缩，扩展业务不中断，减少用户投资成本。

图表16： 华为 GaussDB 300 集群架构



资料来源：华为官网，华泰证券研究所

数据库产品对比分析

这里我们主要讲华为数据库 GaussDB 与目前市场主流数据库：Oracle、AWS 进行对比。

Oracle 数据库作为传统数据库产品的代表，其继续在传统计算场景占据较大的市场份额。同时，针对新兴的计算场景和新的数据类型要求，Oracle 数据库也在进行转型和迭代。

AWS 数据库作为新兴数据库厂商的代表，其数据库产品更多的是为了满足云计算（互联网）场景下对数据处理的需求。目前，AWS 的产品既包括关系型数据库产品，也包括了非关系型数据库产品。

我们这一部分内容，主要介绍下 Oracle 和 AWS 数据库的最新进展。以此来对比下目前华为的 Gauss 数据库与这两大巨头数据库产品之间的异同。

Oracle 数据库

Oracle 作为全球最大、历史最悠久的数据库厂商，在数据库方面的技术领先型和创新性方面，依然保持了强大的优势。Oracle 数据库自从 1978 年发布第一个版本之后，至今已经有 41 年的时间。Oracle 在全球数据库的市场份额超过 50%。其不仅包括商用版的 Oracle DB，同时 MySQL 作为开源数据库，也是 Oracle 旗下的产品。

图表17： Oracle 数据库不断演进

Oracle DB 不断演进



资料来源：Oracle 官网，华泰证券研究所

1、Oracle DB 最新进展

2019 年 Oracle 发布 Oracle DB 19C。在 2019 年 8 月 Oracle 云数据库大会上，Oracle 介绍了其最新产品 Oracle DB 19C 的特性。

Oracle DB 19C 可以在任何环境运行，本次版本增加的新功能包括：

1) 安全可靠的 Oracle 数据库版本

长期支持 12c 第 2 版系列（12.2、18、19）

标准支持服务持续至 2023 年 3 月，扩展支持服务持续至 2026 年 3 月

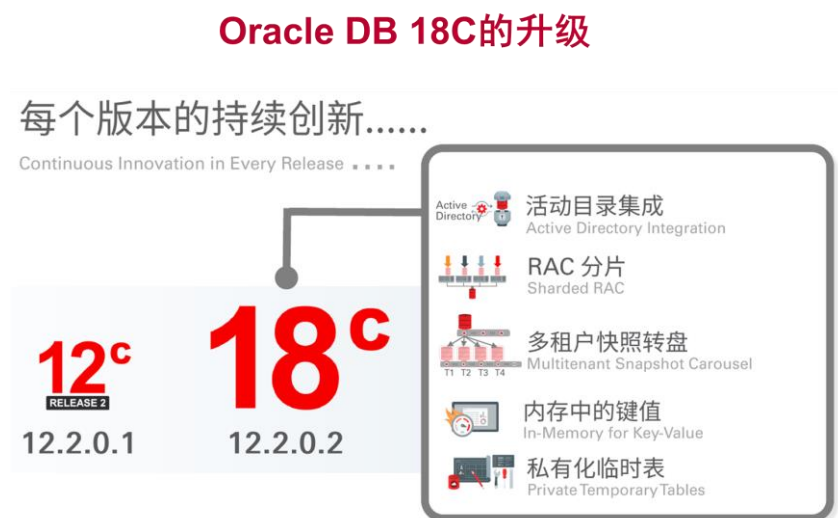
2) 提供更多自治数据库功能

自动索引、实时统计……

3) 提供功能变更以改善整体客户体验

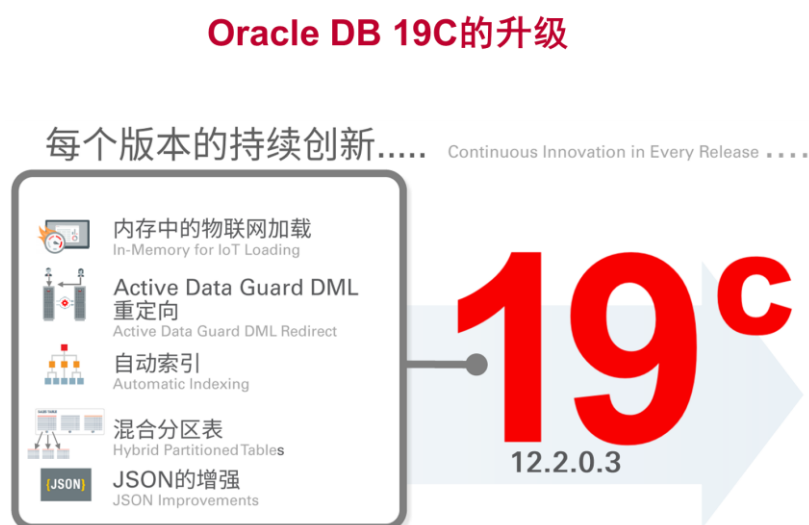
面向物联网的流插入, Active Data Guard 备用数据库更新、针对云对象存储的 SQL 查询。

图表18: Oracle 数据库 18C 的升级



资料来源: Oracle 官网, 华泰证券研究所

图表19: Oracle 数据库 19C 的升级



资料来源: Oracle 官网, 华泰证券研究所

针对以上这些, 我们初看起来, 可能还没有太多直观和概念的认识。我们先看下 Oracle 数据库正在做的数据库革新。主要体现在几个方面:

- 1) 自治数据库 Autonomous Database
- 2) 云 Cloud
- 3) Exadata
- 4) 分片 Sharding
- 5) 区块链 Blockchain
- 6) 微服务 Microservices

我们先看下, Oracle 数据库在自治方面的进展。首先对于“自治 (Autonomous)”, 我们可以理解为将 AI 技术逐步植入和迭加到已有的数据库产品。

Oracle 数据库从 9i 版本，已经开始实现“自动 Autonomous”的功能，包括自动查询重写和自动撤销管理等功能。在之后的 10g、11g、12c 等多个版本中，不断加大和强化 AI 和机器学习技术的应用。

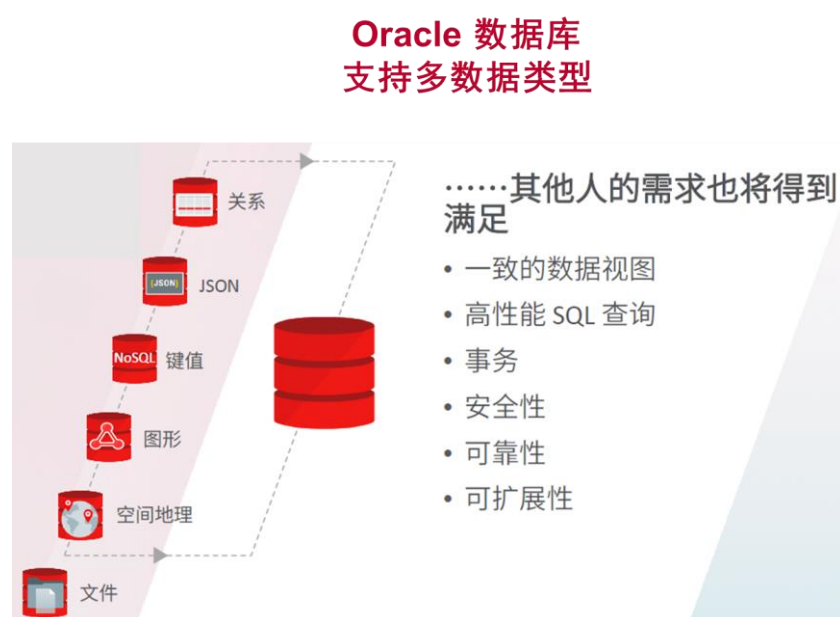
Oracle 的自治数据库，集成了基于机器学习和人工智能技术的监视、管理和分析功能。自治数据库的机器学习和人工智能算法所使用的技术应包括查询优化、自动内存管理和自动存储管理，以提供一个完全自主调优的数据库。自治数据库的目标是，实现数据库调优自动化，防止应用中断，并加强整个数据库应用的安全性。

机器学习算法通过分析大量的记录数据，并标记异常值和异常模式，帮助企业提高数据库安全性，防止入侵者造成破坏。机器学习还可以在系统运行的同时自动、连续地对系统进行打补丁、调整、备份和升级，而无需人工干预。这将尽可能减少发生人为错误或恶意行为的可能性，避免影响数据库运作或安全性。

对于 Oracle 数据库，通常有这样的看法：“Oracle 是一个关系型数据库，因此无法管理和分析非结构化数据”。对于以上这种看法，我们认为，对 Oracle 数据库的认知，还停留在关系型数据库本身的固定范式。

Oracle 是一个多模型数据库，可以支持多个计算场景下的数据管理，包括：文件、图形、键值、空间地理等。因此，Oracle 数据库可以对任何数据类型执行 SQL 查询。

图表20： Oracle 数据库支持多种数据类型



资料来源：Oracle 官网，华泰证券研究所

2、Oracle 数据库的转型

为了应对新的计算环境，新的数据格式以及新的应用场景。目前 Oracle 数据库也在逐步转型。

1) 从单数据库到分布式数据库

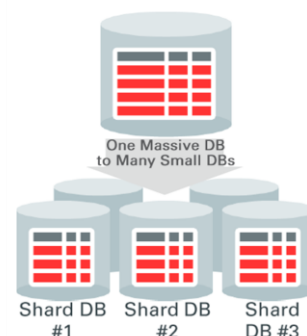
Oracle 的数据库支持的数据计算量都比较大，为了满足全球规模的 OLTP 应用程序对小型数据库的需求，Oracle 利用原生分片技术，将大型数据库分成很多小型数据库。

Oracle 的原生 SQL 对分片表可以做到 1000 个分片
线性扩展的数据和用户隔离到不同的分片
在线添加和重组分片
路由基于分片键，交叉分片查询

图表21： Oracle 的分布式数据库

Oracle数据库原生分片技术

**分片的好处是
可以使用成熟的SQL数据库**



资料来源：Oracle 官网，华泰证券研究所

2) 从关系型到以 Web 为中心的数据

我们之前 Oracle 数据库有分歧的地方在于, Oracle 的数据库不能很好的支持互联网数据。目前, Oracle 数据库逐步向支持互联网场景的数据。

Oracle 利用原生 SQL 对 JSON 进行支持。

- 1) JSON 是 Web 新应用程度最流行的数据格式。
- 2) 数据库中的 JSON 极大简化了应用程序开发
- 3) Oracle 将 JSON 存储在表列中，并支持原生 SQL
- 4) 所有的 Oracle 功能都支持 JSON。

3) 从碎片化数据库到融合数据库

目前，初创格式和云供应商都在大力倡导单用途数据库，其可以避免建立通用数据库做需要的数十年的投资。能在某一个的应用场景快速超越已有的市场领先者。比如，我们现在看到的 AWS 的数据库，其开发了支持不用数据类型和计算场景下的多样化的数据库。

针对市场的多样化数据场景需求和竞争对手的竞争，Oracle 开发了融合数据库，将众多单独的数据库产品融合到 Oracle 数据库产品中，可以支持键值、分析、JSON 等多样化的产品功能需求。

图表22: Oracle 数据库支持多种数据类型



资料来源: Oracle 官网, 华泰证券研究所

AWS 数据库

AWS 数据库的研发和推出, 完全是适应了云应用程序和云计算场景下对数据库新的要求。云时代应用程序需要数据库来存储 TB 到 PB 级的新类型数据, 提供对数据的访问 (毫秒级延迟), 每秒处理数百万个请求, 并扩展以支持位于世界上任何地方的数百万用户。为了满足这些要求, 同时需要关系数据库和非关系数据库作为支撑。

AWS 全托管的数据库服务, 包括: 关系数据库 (适用于事务性应用程序)、非关系数据库 (适用于 Internet 规模应用程序)、数据仓库 (适用于分析)、内存数据存储 (适用于缓存和实时工作负载)、图形数据库 (适用于构建具有高度互联数据的应用程序)、时间序列数据库 (适用于衡量随时间的变化) 和分类账数据库 (用于维护完整且可验证的交易记录)。

1、AWS 关系型数据库: Aurora

关系数据库存储具有预定义架构的数据及其之间的关系, 设计用于支持 ACID 事务、维护引用完整性和数据一致性。

应用场景: 传统应用程序、ERP、CRM 和电子商务

AWS 产品:

Amazon Aurora
MySQL、PostgreSQL
Amazon RDS
MySQL、PostgreSQL、MariaDB、Oracle、SQL Server
Amazon Redshift

2、键值数据库: DynamoDB

键值数据库经过优化, 以毫秒级的速度存储和检索大量键值对, 避免关系数据库的性能开销和规模限制。

应用场景: Internet 规模的应用程序、实时竞价、购物车和客户喜好

AWS 产品:
Amazon DynamoDB

3、文档数据库：DocumentDB

文档数据库设计用于存储文档等半结构化的数据，可供开发人员直观地使用，因为数据通常以可读文档的形式呈现。

应用场景：内容管理、个性化和移动应用程序

AWS 产品：

Amazon DocumentDB（兼容 MongoDB）

4、内存数据库：

内存中数据库用于需要实时访问数据的应用程序。通过直接将数据存储在内存中，这些数据库为不满足于毫秒级延迟的应用程序提供微秒级延迟。

应用场景：缓存、游戏排行榜和实时分析

AWS 产品：

Amazon ElastiCache for Redis

适用于 Memcached 的 Amazon ElastiCache

5、图形数据库

对于需要让数百万用户查询和浏览高度互联的图形数据集之间的关系并实现毫秒级延迟的应用，可以使用图形数据库。

应用场景：欺诈检测、社交网络和建议引擎

AWS 产品：

Amazon Neptune

6、时间序列数据库

时间序列数据库用于从大量随时间变化的数据（称为时间序列数据）高效地收集、合并和提炼见解。

应用场景：IoT 应用、开发运营和工业遥测

AWS 产品：

Amazon Timestream

7、分类账数据库

在您需要集中的可信权威方式来维护可扩展、完整、加密且可验证的事务记录时，可以使用分类账数据库。

应用场景：系统记录、供应链、注册和银行事务。

AWS 产品：

Amazon Quantum Ledger Database (QLDB)

Amazon Aurora 详细介绍

Amazon Aurora 是一种与 MySQL 和 PostgreSQL 兼容的关系数据库，专为云而打造，既具有传统企业数据库的性能和可用性，又具有开源数据库的简单性和成本效益。

Amazon Aurora 的速度最高可以达到标准 MySQL 数据库的五倍、标准 PostgreSQL 数据库的三倍。它可以实现商用数据库的安全性、可用性和可靠性，而成本只有商用数据库的 1/10。Amazon Aurora 由 Amazon Relational Database Service (RDS) 完全托管，RDS 可以自动执行各种耗时的管理任务，例如硬件预置以及数据库设置、修补和备份。

Amazon Aurora 采用一种有容错能力并且可以自我修复的分布式存储系统，这一系统可以把每个数据库实例扩展到最高 64TB。它具备高性能和高可用性，支持最多 15 个低延迟读取副本、时间点恢复、持续备份到 Amazon S3，还支持跨三个可用区 (AZ) 复制。

图表23： Amazon Aurora 典型客户



资料来源：AWS 官网，华泰证券研究所

Amazon DynamoDB 详细介绍

Amazon DynamoDB 是一个键/值和文档数据库，可以在任何规模的环境中提供个位数的毫秒级性能。它是一个完全托管、多区域多主的持久数据库，具有适用于 Internet 规模的应用程序的内置安全性、备份和恢复和内存缓存。DynamoDB 每天可处理超过 10 万亿个请求，并可支持每秒超过 2000 万个请求的峰值。

许多全球发展最快的企业，如 Lyft、Airbnb 和 Redfin，以及 Samsung、Toyota 和 Capital One 等企业，都依靠 DynamoDB 的规模和性能来支持其关键任务工作负载。

数十万 AWS 客户选择 DynamoDB 作为键值和文档数据库，用于其移动、Web、游戏、广告技术、物联网以及其他需要任何规模的低延迟数据访问的应用程序。

图表24： Amazon DynamoDB 典型客户

AWS Dynamo典型客户

<p>Nike</p> <p>Nike Digital 将其大型 Cassandra 集群迁移到完全托管式 Amazon DynamoDB，从而留出更多资源，打造更好客户体验。</p>	<p>GE Aviation</p> <p>GE Aviation 使用 Amazon DynamoDB 重新构建了其测绘和数据查询应用程序，实现了成本节约、可扩展性和优越性能。</p>
<p>Samsung</p> <p>Samsung Electronics 使用 DynamoDB 来执行 PB 级移动应用程序备份，实现了一致的高性能和成本节约。</p>	<p>Netflix</p> <p>Netflix 使用 DynamoDB 来运行 A/B 测试，为 1.25 亿客户打造了个性化的流式体验。</p>
<p>Lyft</p> <p>Lyft 利用 DynamoDB 的可扩展性实现了多数据存储，包括为所有乘车服务存储 GPS 坐标的乘车跟踪系统。</p>	<p>Snap</p> <p>Snap 将最大的存储工作负载“Snapchat Stories”迁移了 DynamoDB，在提高了性能的同时降低了成本。</p>

资料来源：AWS 官网，华泰证券研究所

华为 GaussDB：对比

我们以华为 GaussDB 200 为例，来看下目前 GaussDB 的下游应用情况。

2018 年，GaussDB 200 融合数据仓库荣获 2018 数博会领先科技成果奖。2017~2019 年连续 3 年入围 Gartner “Magic Quadrant for Data Management Solutions for Analytics”，中国区厂商排名第一。

应用场景：华为 GaussDB 200 分布式 OLAP 数据库已经赢得全球 300 多个客户的信赖，广泛应用于政府、金融、公共安全、运营商、大企业等行业。

1) 金融领域

银行融合数仓，多类型数据统一存储和分析，性能大幅提升。

保险营销管理平台，提升服务水平，显著提升续保率，交叉拓展率等指标，客户流失率降幅高达 50% 以上。

2) 电子政务

税务大数据，提升纳税人办税效率，提升工作人员办公效率，推动整体涉税服务迭代升级。

财务云大数据，打破财政应用系统条块分割，以统一规范的数据标准和数据结构为基础，预算编制、绩效监督、综合管理等财政所有业务上云。

警务大数据，“两抢一盗”案件率明显下降。

3) 运营商

集中经分，利用 GaussDB 多租户能力对资源进行集中管控，支撑客户群分析从 10 分钟缩短到实时分析。

详单查询，日增 600GB 数据持续入库，查询时间从分钟级减少为平均 5s。

图表25： 华为 GaussDB 200 典型客户

华为 GaussDB 200 典型客户



资料来源：华为官网，华泰证券研究所

对比一：产品矩阵对比

通过以上三家数据库产品对比，我们可以看出：

- 1) Oracle 的数据库在关系型数据库领域的地位还是非常稳固，其也在非关系型数据库应用的场景不断迭代自身产品，而使其更加适应更多的计算场景。
- 2) AWS 数据库产品虽然发展比较晚，但产品线还是很齐全，既包括关系型数据库，也包括非关系型数据库。
- 3) 华为 Gauss 数据库已经发布了 GaussDB100、GaussDB200、GaussDB300 三个产品系列。

对比二：应用场景对比

- 1) Oracle 数据库产品 OLTP 和 OLAP 的产品都具备，而且功能比较强大。
- 2) AWS 数据库产品，也同时具备 OLTP 和 OLAP 的产品线。
- 3) 华为 Gauss 数据库产品中，GaussDB 100 更多以 OLTP 为主，GaussDB 200 主要以 OLAP 为主，GaussDB 300 兼顾了 OLTP 和 OLAP。在 OLTP 方面，主要客户为招商银行；OLAP 方面，包括工商银行等。

对比三：客户属性对比

- 1) Oracle 的数据库产品在全球跨国企业、金融领域、能源领域、航空领域等具备强大的竞争优势。
- 2) 我们认为，AWS 的数据库在对新兴计算场景偏重的互联网企业中，更加具备竞争优势。
- 3) 华为 Gauss 数据库已经在国内开始逐步推广，从目前客户类型来看，主要以国内的银行、保险、能源等行业客户为主。从客户范围和数量方面，距离 Oracle 和 AWS 还有一定差距。

对比四：数据库支持的架构

- 1) Oracle 数据库产品主要支持 X86 的架构。我们知道 Oracle 的数据库、Window 操作系统和 Intel 的芯片，在 PC 时代是一套完整的生态体系，它们之间是高度耦合的。产品迭代的时候都会以彼此产品作为基础，不断优化做到高度适配。
- 2) AWS 数据库产品也只支持 X86 架构。
- 3) 华为 Gauss 数据库产品可以支持 X86、ARM、GPU、NPU 等异构计算。可以看出，Gauss 数据库未来的目标市场不只是国产服务器市场，可能还包括未来 5G 带来更多的计算应用场景。

总结

通过以上我们对华为 Gauss 数据库的分析，以及与其他产品的对比，我们对 Gauss 数据库应该有了一个比较立体的认知和理解。最后，将我们的观点总结如下。

1、从发展时间来看，华为的数据库从起步到现在已经有 12 年的时间，目前正处于快速迭代和应用期。从 Oracle 数据库和 AWS 数据库发展来看，Oracle 数据库从 1978 年开始到现在已经有 41 年的历史，AWS 的数据库从 2006 年开始到现在大概走过了 13 个年头，而 AWS 也算是数据库领域的新进入者。因此对于华为 Gauss 数据库而言，未来的路还很长。

2、从生态体系来看，Oracle 的数据库成为全球第一数据库的地位，也是经过了很多次版本的更新升级，更与上世纪 80 年代开始的全球 IT 生态体系的逐步确立有关。Oracle 数据库世界霸主地位，是随着 Windows 操作系统、Intel X86 芯片一起建立的 PC 时代的 IT 底层生态而逐步确立的。而 AWS 数据库则是适应了互联网时代（云计算时代）新的计算场景对数据库的新需求，再加上自身的云生态体系，逐步迎来的客户使用的推广。从华为 Gauss 数据库来看，华为 IT 架构的底层生态已经逐步建立起来，包括芯片、操作系统、数据库等，这些在华为 IT 体系内部是高度耦合的。

3、从计算场景来看，数据库也有其使用的计算场景。我们通过 Oracle 数据库、AWS 数据库，包括一些新兴的数据库公司，从他们产品的和客户使用来看，IT 发展的不同阶段，在不同的计算场景、不同的数据类型环境下，对数据库的要求是不同的。从而带来一个问题，数据库是否也有其生命周期？当无论答案是或否，但其内在本质是，需要数据库产品不断根据外界变化，而进行更新升级，甚至自我革新。Oracle 的数据库诠释了 PC 时代的霸主地位，AWS 数据库演绎了互联网云时代的新兴角色，华为的 Gauss 数据库是否能在未来计算环境的变化中迎来新的产业地位？

4、产品的优化和升级是一条漫漫长路。这也是软件产品的一个特点，所有我们感觉好用的软件产品都不是短期开发出来的，而是靠时间堆出来的，都是经过无数客户的实用反馈，反复修改 bug，优化到大部分用户都满意的水平。对于数据库产品更是如此，Oracle 的数据库之所以好用在于其在产品迭代中，不断收到客户的使用反馈，不断地对产品进行相应的修改和更新，这一过程进行了 41 年。对于华为 Gauss 数据库产品也是如此，不过我们现在已经看到，Gauss 数据库已经在跟银行大客户合作，这对于华为而言，是其数据库快速发展的一个重要因素。

风险提示

- 1、数据库技术升级速度低于预期的风险；
- 2、数据库产品应用推广低于预期的风险；
- 3、数据库应用场景拓展低于预期的风险。

免责声明

本报告仅供华泰证券股份有限公司（以下简称“本公司”）客户使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司对该等信息的准确性及完整性不作任何保证。本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，本公司可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。本公司不保证本报告所含信息保持在最新状态。本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成所述证券的买卖出价或征价。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，本公司及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本公司及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，也可能为之提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权力。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

本公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：91320000704041011J。

全资子公司华泰金融控股（香港）有限公司具有香港证监会核准的“就证券提供意见”业务资格，经营许可证编号为：A0K809

©版权所有 2019 年华泰证券股份有限公司

评级说明

行业评级体系

一报告发布日后的 6 个月内的行业涨跌幅相对同期的沪深 300 指数的涨跌幅为基准；

一投资建议的评级标准

增持行业股票指数超越基准

中性行业股票指数基本与基准持平

减持行业股票指数明显弱于基准

公司评级体系

一报告发布日后的 6 个月内的公司涨跌幅相对同期的沪深 300 指数的涨跌幅为基准；

一投资建议的评级标准

买入股价超越基准 20%以上

增持股价超越基准 5%-20%

中性股价相对基准波动在-5%~5%之间

减持股价弱于基准 5%-20%

卖出股价弱于基准 20%以上

华泰证券研究

南京

南京市建邺区江东中路 228 号华泰证券广场 1 号楼/邮政编码：210019

电话：86 25 83389999/传真：86 25 83387521

电子邮件：ht-rd@htsc.com

深圳

深圳市福田区益田路 5999 号基金大厦 10 楼/邮政编码：518017

电话：86 755 82493932/传真：86 755 82492062

电子邮件：ht-rd@htsc.com

北京

北京市西城区太平桥大街丰盛胡同 28 号太平洋保险大厦 A 座 18 层

邮政编码：100032

电话：86 10 63211166/传真：86 10 63211275

电子邮件：ht-rd@htsc.com

上海

上海市浦东新区东方路 18 号保利广场 E 栋 23 楼/邮政编码：200120

电话：86 21 28972098/传真：86 21 28972068

电子邮件：ht-rd@htsc.com