- 单表查询的常见执行计划有：全表扫描（ TABLE FULL SCAN ）、索引扫描（TABLE RANGE SCAN）、主键唯一查询（TABLE GET）

- OceanBase 数据库通过 RootService 提供智能的负载均衡能力，在水平扩缩容和数据动态变化场景下可以达到各个服务节点上分区数与存储容量的均衡，以及不同Zone间Leader副本的均衡。(日志副本而非tablet)

租户类型

- 系统租户是集群默认创建的租户，与集群的生命周期一致，负责管理集群和所有租户的生命周期。系统租户的数据是集群私有的，不支持主备集群物理同步和物理备份恢复。

- 用户租户是由用户创建的租户，对外提供完整的数据库功能，支持 MySQL 和 Oracle 两种兼容模式。支持跨集群物理同步和物理备份恢复，典型数据包括：Schema 数据、用户表数据及事务数据等。

- Meta 租户是 OceanBase 数据库内部自管理的租户，每创建一个用户租户系统就会自动创建一个对应的 Meta 租户，其生命周期与用户租户保持一致。用于存储和管理用户租户的集群私有数据，这部分数据不需要进行跨库物理同步以及物理备份恢复，这些数据包括：配置项、位置信息、副本信息、日志流状态、备份恢复相关信息、合并信息等。

- 日志流的成员：leader-follower，仲裁-普通，全功能-只读

- LOAD DATA从OSS或登录到OBServier中运行，不能通过OBProxy

- LS 数量 = UNIT\_NUM \* first\_level\_primary\_zone\_num)

- 参与 Paxos 多数派同步的副本拥有完整的日志和数据，也可以参与 Leader 选举或者当选为 Leader，这类副本叫全功能副本。全功能副本<zone\_num

- Zone 是一个逻辑概念，一个 Zone 内的所有OBServer共同承载一份完整的副本。

- 事务日志需要在多数派个副本同步成功才算提交完成。

- OBClient同时兼容访问OceanBase数据库的MySQL以及Oracle租户

- OceanBase 数据库社区版仅提供了MySQL 模式

- OceanBase 企业版和社区版都可以进行单机集中式部署和集群高可用部署。

- 安装完OCP之后,可以利用OCP单独部署OBProxy

- OBD是OceanBase 集群安装部署工具，通过命令行部署或白屏界面部署的方式

- OAT(OceanBase 管理者工具)是一个可视化平台，用于安装和管理OceanBase 生态产品和组件。

- 通过OCP可对主机进行：安装OCP Agent、添加主机、删除主机

OCP软件目录：

- OCP程序：/home/admin/ocp-server/

- OCP主进程日志：/home/admin/logs/ocp/

- OCP任务日志：/home/admin/logs/task/

- OCP文件本地缓存：/home/admin/data/files/

- OCP Agent包括进程：ocp\_agentd、ocp\_mgragent、ocp\_monagent

- 通过OCP可对租户：创建主租户 、调整租户资源、锁定租户、创建备租户

- OBClient 是一个交互式和批处理查询工具，需要单独安装。它是一个命令行用户界面

OceanBase扩容步骤（添加zone和observer）

- 执行alter system add server;命令，将启动的observer添加到集群

- 执行alter resource pool unit\_num=;命令，扩充资源池中的unit个数

- 在扩容服务器上，正确启动observer进程

- 租户扩容,可先通过添加服务节点,完成集群扩容,再通过增加资源单元的个数完成租户扩容

- observer进程异常休止的时长大于等于该参数设置的值,OceanBase数据库会将机器做"临时下线"处理,此时不会为故障节点上所保存的所有分区进行补副本操作；observer进程异常终止的时长小于该参数设置的值.由于节点故障时间较短认为此进程有可能较快恢复,因此OceanBase数据库暂时不做处理,以避免频繁的数据迁移

- 两阶段提交协议中，OBServer是协调者

- OB Proxy需要部署到一台独立的服务器上， 以保证其性能要求

- OceanBase 支持的事务隔离：RC、Serializable、RR

- 假设OceanBase有3个Zone，其中2个Zone部署在一个城市的两个机房中，另外一个Zone部署在另外一个城市的一个机房中。如果同城的一个机房宕机，则会增加异地同步延迟。

- 建立 table group 的主要目的是：减少跨机分布式事务

- OceanBase允许你通过设置租户的 primary\_zone 属性来精确控制主副本的分布偏好。可以打散或聚焦。

- MetaDB：基于容器部署的 OceanBase 数据库服务，可用于OCP、ODC 和 OMS 等产品的元数据存储。

- 不同的Zone可以对应不同城市、一个城市的不同机房、或者一个机房的不同机架。

- 在 OceanBase V4， 同一个日志流中的不同分区的 Leader 副本可以分布在不同的 OBServer 上。 （ 错误 ）

- 资源池是资源分配的基本单位，同一个资源池内的各个Unit具有相同的资源规格，即该资源池内Unit 的物理资源大小都相同。一个资源单元只能属于一个租户；一个租户在一个OBServer上只能同时存在一个资源单元。

- OceanBase 有自动的负载均衡能力，可以将Table、Partition均匀地分散在Zone内的多个OBServer节点上，将Leader副本均匀地分散在指定的多个PrimaryZone中

索引类型

- 全局索引：全局索引是在整个表中创建的索引，可以包含表的所有分区的数据。全局索引适用于跨分区进行查询的场景。

- 局部索引：局部索引是在每个分区中创建的索引，仅包含该分区的数据。局部索引适用于在分区内部进行查询的场景。

- 空间索引：主要应用于地理信息系统（GIS）和空间数据处理领域

- MySQL 模式默认创建的是本地索引（local），Oracle 模式默认创建的是全局索引（global）。

与局部索引不同，全局索引不再和主表的分区保持一对一的关系，而是将所有主表分区的数据合成一个整体来建立全局索引。全局唯一索引可以保证表数据的全局唯一性。

LSMTree

优点：通过追加写的方式，解决了随机写和写放大的问题，适合于大量数据更新的场景。

缺点：读取数据的路径长，读取的性能较传统的存储引擎差。

转储合并

- 动态数据：MEMTABLE，LSMTree 架构中增量数据在写内存中的组织形式。插入/更新/删除数据时，数据被写入内存块，在 HashTable 和 BTree 中存储的均为指向对应数据的指针

- 静态数据：SSTABLE，LSMTree 架构中表数据在磁盘中的存储形式，由数据宏块和微块组成。

- 数据落盘（转储）：当 MemTable 的大小超过一定阈值时，就需要将 MemTable 中的数据转存到 SSTable 中以释放内存。采用顺序追加写入的方式，避免随机写和写放大，提升写入的性能。

- 第一次压缩是编码压缩

- 第二次压缩是通用压缩

- OceanBase 采用分层转储的策略，由系统后台任务自动完成增量数据的压缩与合并，提升了转储的效率和数据读取的性能。

L0层：从内存中转储生成的SSTable，称为L0层

L1层：当L0层的数据量达到阈值时，会自动将L0层的多个SSTable压缩成一个SSTable，称为L1层

L2层：将所有的动静态数据做归并，形成一个新的基线SSTable，称为L2层

- MVCC：多版本并发控制。使得读写操作可以基于不同版本的数据进行， OceanBase通过锁与多版本并发控制（MVCC）机制⼀起来控制并发的读、写操作，实现了读不加锁，避免读写冲突

- OceanBase不支持读未提交。仅MySQL模式支持可重复读。支持读已提交（Read Committed） 可串行化（Serializable）

- Paxos支持乱序的日志同步，支持事务的高并发执行。

- 在OB里，只要进入commit阶段，事务一定会提交

- OceanBase 事务引擎具备强大的分布式事务处理能力，保证事务的原子性、一致性、隔离性和持久性（ACID）。

- 全局时间戳服务：通过全局时间戳服务，保证全局的事务顺序和逻辑一致。

- 两阶段提交协议：通过优化的两阶段提交协议，确保所有事务的参与节点都能成功提交或回滚事务。

- 一致性算法：通过Paxos一致性算法确保分布式环境中的一致性。

- 集群高可用部署保证RPO=0

连接方式

- mysql：命令行客户端

- OBClient：命令行用户界面，支持mysql和oracle租户

- ODC：开发者中心

驱动程序

- Java: MySQL Connector/J（MySQL官方提供JDBC驱动）、OceanBase Connector（支持MySQL和Oracle）

- C: OceanBase Connector/C

- Python: PyMySQL

- 租户创建后类型无法修改

- 创建索引要遵循以下原则：

- 覆盖索引：尽可能利用包含所有必要数据的索引，在满足查询需求的前提下，可以避免访问表数据。

- 等值优先：对于组合索引，建议将等值条件的列放在最前面。

- OceanBase数据库中主键:非空且唯一,全局唯一,数量唯一,主键即表(数据按照主键顺序进行存储，并不会为主键单独创建一个索引。)

- 在OceanBase数据库中，唯一索引:允许空值且唯一，可以包含重复的NULL值记录。非NULL键值在全表范围内唯一。全局唯一,约束即索引,索引键包含分区键S\_ID，因此分区唯一即全局唯一，可以创建为局部索引。

- Ob支持的联接

内联接INNER JOIN、左（外）联接LEFT [OUTER] JOIN、右（外）联接RIGHT [OUTER] JOIN、全（外）联接FULL [OUTER] JOIN

- GV$OB\_SQL\_AUDIT：全局SQL 审计表，可以用来查看每次请求客户端来源、执行服务器信息、执行状态信息、等待时间以及执行各阶段耗时等

- EXPLAIN命令的语法：EXPLAIN [BASIC | EXTENDED ] explainable\_stmt;

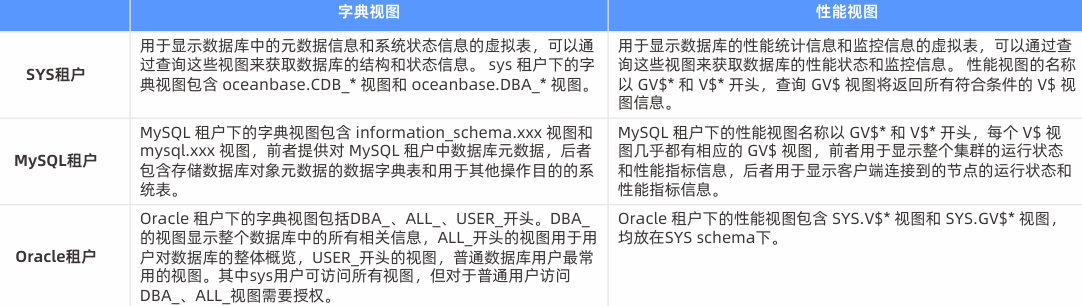
- EXPLAIN BASIC：用于最基本的计划展示，例如算子名称、所访问的表和索引名，以及过滤信息（Filter）等。BASIC是默认的输出模式。

- EXPLAIN EXTENDED：用于最详细的计划展示，可以帮助你理解优化器的执行计划选择，从而解决SQL执行计划问题。

- 一条语句只能包含一个 Hint 注释，并且该注释必须跟随 SELECT、UPDATE、INSERT、REPLACE 或 DELETE 关键字

- 数据导出方式包括：ODC导出、OBDUMPER、OUTFILE

- 数据导入方式包括：ODC导入、OBLOADER、LOAD DATA（包括常规导入和旁路导入）、外表



- 日志文件：文件分为启动运行日志observer.log、选举模块日志election.log 和 rootservice.log 三种类型。

- 日志级别：分为ERROR、WARN、INFO、EDIAG、WDIAG、TRACE、DEBUG。

- 配置项（PARAMETER）

- 集群级配置项：用于整个OceanBase数据库集群的配置选项，通常在集群启动时进行配置，配置后不轻易修改

- 租户级配置项：指适用于租户级别的配置选项，租户创建和管理时进行配置，随时修改

- 系统变量（VARIABLE）

- Global 变量：同一租户内的不同用户共享全局变量。修改不会随会话的退出而失效。此外，全局变量修改后，需要重新建立Session 才能生效。

- Session 变量：当客户端连接到数据库后，数据库会复制全局变量来自动生成Session 变量

- 生效范围：配置项分为集群、Zone、机器（某台 OBServer）和租户；系统变量分为租户的Global或Session级别。

- 生效方式：配置项分为动态和重启生效；Session系统变量只对当前会话有效，Global级系统变量对新会话有效

- 事务类型:普通事务即单机事务、分布式事务即多机事务、XA事务特指跨数据库集群或租户的分布式事务。

- 事务状态：长事务（执行时间超过 60s 不提交）、悬挂事务（进入提交阶段，但无法成功）、其他事务

