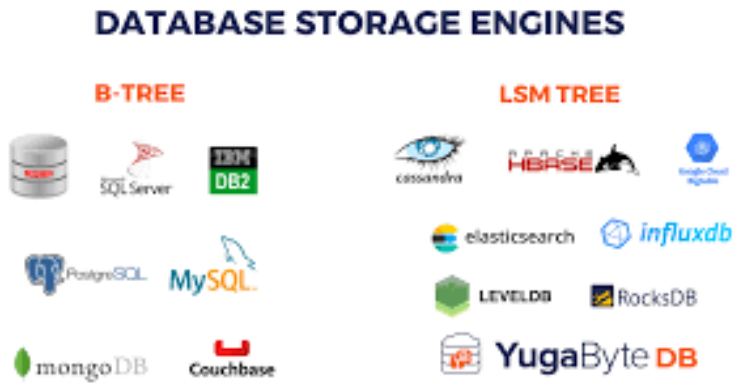
MySQL事务和存储引擎



LSM: the algorithmic limit

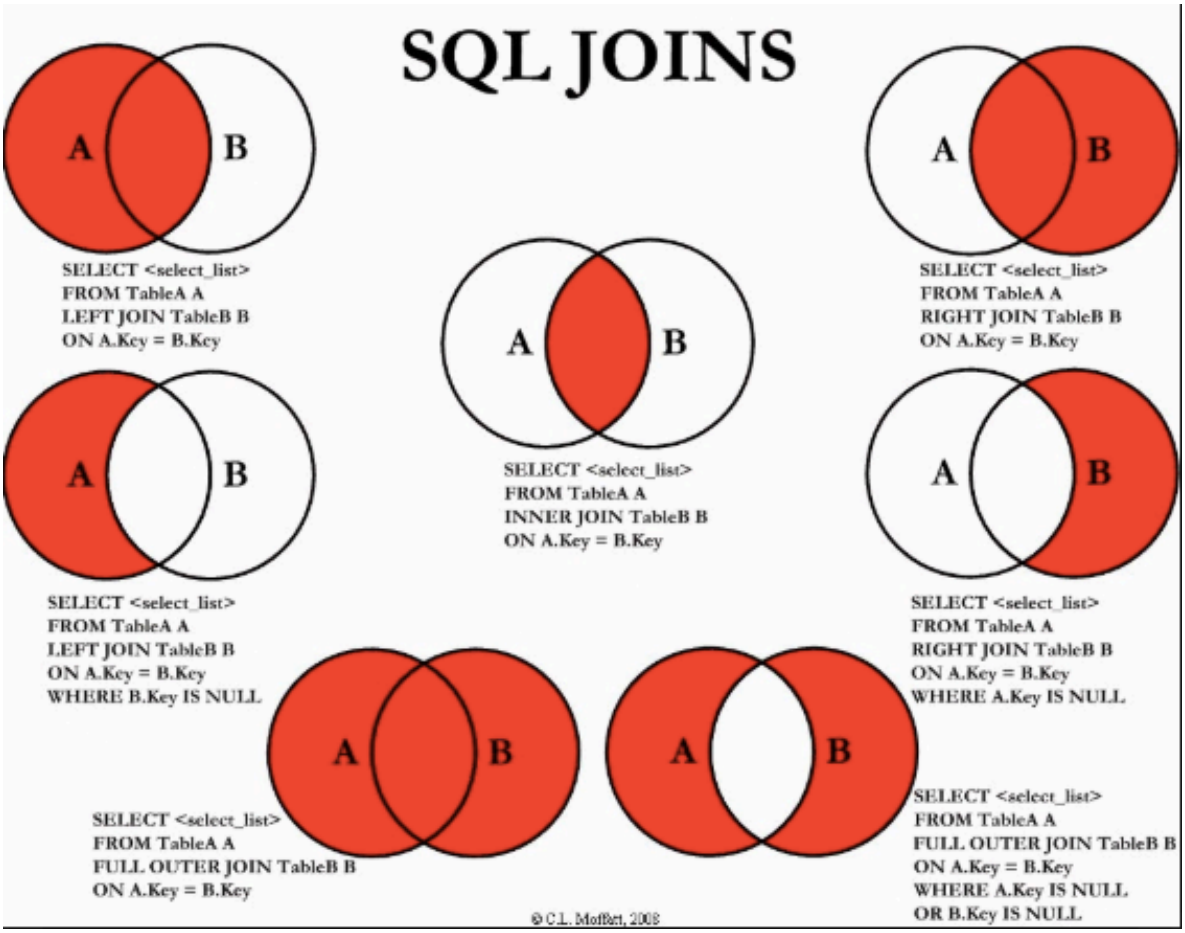
	LSM tree	B-tree
Search	K * $O(\log_2 N/B)$	$O(\log_B N)$
Delete	$O(\log_2(N)/B)$	$O(\log_B N)$
Insert	$O(\log_2(N)/B)$	$O(\log_B N)$

Employee table

LastName	DepartmentID
Rafferty	31
Jones	33
Heisenberg	33
Robinson	34
Smith	34
Williams	NULL

Department table

DepartmentID	DepartmentName
31	Sales
33	Engineering
34	Clerical
35	Marketing



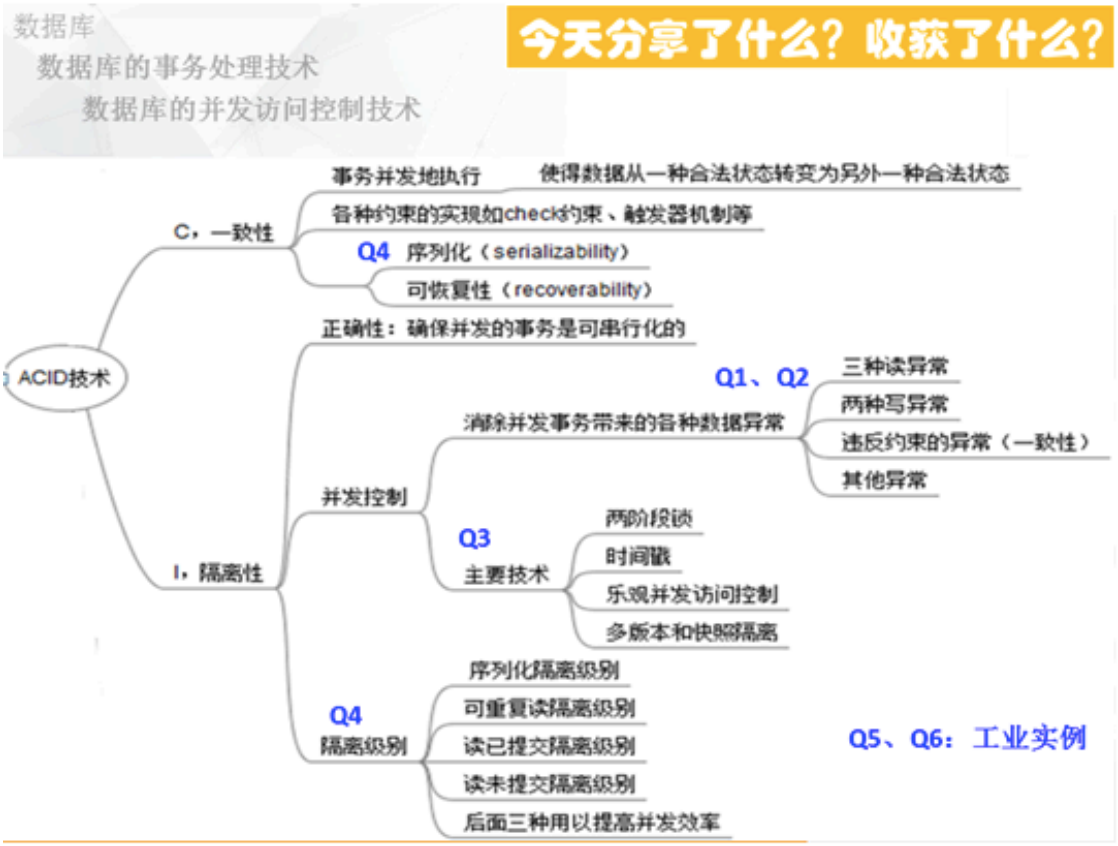
http://www.vldb.org/conf/1990/P186.PDF
https://en.wikipedia.org/wiki/Category:Join_algorithms

Join Type	Memory (blocks)	CPU	I/O	Temp. Tables
Simple	2	$o(n^2)$	b^2	no
Nested Loop	b	$o(n^2)$	2b	no
Sort/Merge	3	o(n log n)	6b+4b log b	yes
	$\sqrt{b}$	o(n log n)	8b	yes
	b	o(n log n)	4b	yes
	2b	o(n log n)	2b	no
Simple Hash	b	o(n)	2b	no
Grace Hash	$\sqrt{b}$	o(n)	6b	yes
Hybrid Hash	$\sqrt{b}$...b	o(n)	6b ... 2b	yes

Table 1: Cost Functions of Join Algorithms

	比 较 项 目	联 机 业 务	批处理业务
业务特征	操作特点	日常业务操作，尤其是包含大量前台操作	后台操作，例如统计报表、大批量数据加载
	响应速度	优先级最高，要求响应速度非常高	要求速度高、吞吐量大
	吞吐量	小	大
	并发访问量	非常高	不高
	单笔事务的资源消耗	小	大
	SQL 语句类型	主要是插入和修改操作（DML）	主要是大量查询操作或批量 DML 操作
技术运用	索引类型	B*索引	Bitmap、Bitmap Join 索引
	索引量	适量	多
	访问方式	按索引访问	全表扫描
	连接方式	Nested_loop	Hash Join
	BIND 变量	使用或强制使用	不使用
	并行技术	使用不多	大量使用
	分区技术	使用，但目标不同	使用，但目标不同
	物化视图	少量使用	大量使用

李海翔　腾讯　T4
http://www.techweb.com.cn/news/2017-05-15/2523622.shtml



数据库隔离级别：

Table 4. Isolation Types Characterized by Possible Anomalies Allowed.

Isolation level	**P0** Dirty Write	**P1** Dirty Read	**P4C** Cursor Lost Update	**P4** Lost Update	**P2** Fuzzy Read	**P3** Phantom	**A5A** Read Skew	**A5B** Write Skew
READ UNCOMMITTED == Degree 1	Not Possible	Possible	Possible	Possible	Possible	Possible	Possible	Possible
READ COMMITTED == Degree 2	Not Possible	Not Possible	Possible	Possible	Possible	Possible	Possible	Possible
Cursor Stability	Not Possible	Not Possible	Not Possible	Sometimes Possible	Sometimes Possible	Possible	Possible	Sometimes Possible
REPEATABLE READ	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible	Possible	Not Possible	Not Possible
Snapshot	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible	Sometimes Possible	Not Possible	Possible
ANSI SQL SERIALIZABLE == Degree 3 == Repeatable Read Date, IBM, Tandem, ...	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible	Not Possible

精通spring4.x 企业应用开发实战

表 11-1　事务隔离级别对并发问题的解决情况

隔离级别	脏　读	不可重复读	幻　象　读	第一类丢失更新	第二类丢失更新
READ UNCOMMITED	允许	允许	允许	不允许	允许
READ COMMITTED	不允许	允许	允许	不允许	允许
REPEATABLE READ	不允许	不允许	允许	不允许	不允许
SERIALIZABLE	不允许	不允许	不允许	不允许	不允许

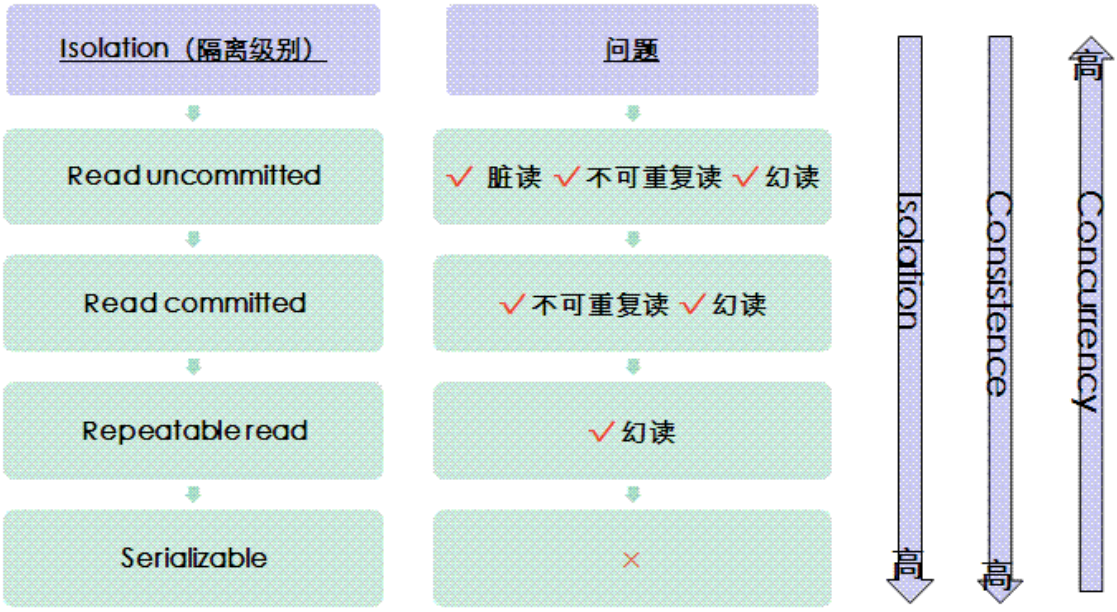
wiki
https://en.wikipedia.org/wiki/Isolation_(database_systems)

Isolation level	Lost updates	Dirty reads	Non-repeatable reads	Phantoms
Read Uncommitted	don't occur	may occur	may occur	may occur
Read Committed	don't occur	don't occur	may occur	may occur
Repeatable Read	don't occur	don't occur	don't occur	may occur
Serializable	don't occur	don't occur	don't occur	don't occur

SQL92

Isolation Level	Dirty Read	Nonrepeatable Read	Phantom Read
Read uncommitted	Possible	Possible	Possible
Read committed	Not possible	Possible	Possible
Repeatable read	Not possible	Not possible	Possible
Serializable	Not possible	Not possible	Not possible

Isolation level	Dirty read	Nonrepeatable read	Phantom
Read uncommitted	Yes	Yes	Yes
Read committed	No	Yes	Yes
Read committed using row versioning	No	Yes	Yes
Repeatable read	No	No	Yes
Snapshot	No	No	No
Serializable	No	No	No



Oracle 支持以下三种事务隔离级别（transaction isolation level）。

隔离级别	描述
——	
已提交读取	Oracle 默认使用的事务隔离级别。事务内执行的查询只能看到查询执行前（而非事务开始前）就已经提交的数据。Oracle 的查询永远不会读取脏数据（未提交的数据）。 Oracle 不会阻止一个事务修改另一事务中的查询正在访问的数据，因此在一个事务内的两个查询的执行间歇期间，数据有可能被其他事务修改。举例来说，如果一个事务内同一查询执行两次，可能会遇到不可重复读取或不存在读取的现象。
串行化	串行化隔离的事务只能看到事务执行前就已经提交的数据，以及事务内 INSERT， UPDATE，及 DELETE 语句对数据的修改。串行化隔离的事务不会出现不可重复读取或不存在读取的现象。
只读模式	只读事务只能看到事务执行前就已经提交的数据，且事务中不能执行 INSERT， UPDATE，及 DELETE 语句。

Application	Isolation
SQLite	Strong isolation
MariaDB	Four isolation levels in SQL standards [38]
Kyoto Cabinet	Intentionally no isolation
APT	No isolation
vim	Strong isolation or no isolation

Table 2: Isolation levels in five real-world applications.

Application	Lines of code		Mechanisms to guarantee application consistency	
	Original	Modified	Original	Modified
SQLite	217,313	38	Physical RBJ or WAL	Specifying an atomic code region
MariaDB	1,534,980	240	Physiological WAL & double-write	Physiological WAL & atomic database write
Kyoto Cabinet	162,606	26	Physical RBJ on a `mmap`-ed region	Specifying an atomic code region
APT	407,642	4	None	Specifying an atomic code region
vim	1,179,246	9	`rename`-based update	Specifying an atomic code region
Total	**3,501,787**	**317**		

NOTE. RBJ (rollback journaling), WAL (write-ahead-logging)

man
info

cd /usr/local/mysql/bin/
sudo /usr/local/mysql/support-files/mysql.server start

https://www.jb51.net/article/109348.htm
mysqladmin -h localhost -u root -p ping
mysqladmin -h localhost -u root -p status
mysqladmin -h localhost -u root -p extended-status
mysqladmin -h localhost -u root -p variables

mysql -uroot -p
密码 1234

A:redo
C:undo

I:lock:行锁，表锁，意向锁，record-lock，gap-lock，next-key lock
D:redo

```
CREATE PROCEDURE p_load (count int unsigned)
BEGIN
declare s int unsigned default 1;
declare c char(80) default repeat('a',80);
while s <= count do
insert into test_load select null,c;
commit;
set s = s + 1;
end while;
END
```

sysbench --test=fileio help

https://www.percona.com/live/europe-amsterdam-2015/sites/default/files/slides/MySQLOptimizerOverview2015.pdf

root@localhost: q7670ykW:i0J
If you lose this password, please consult the section How to Reset the Root Password in the MySQL reference manual.

mysql -h rm-tatbb8.mysql.rdstest.tbsite.net -P 3306 -ubb8 -pbb8124309

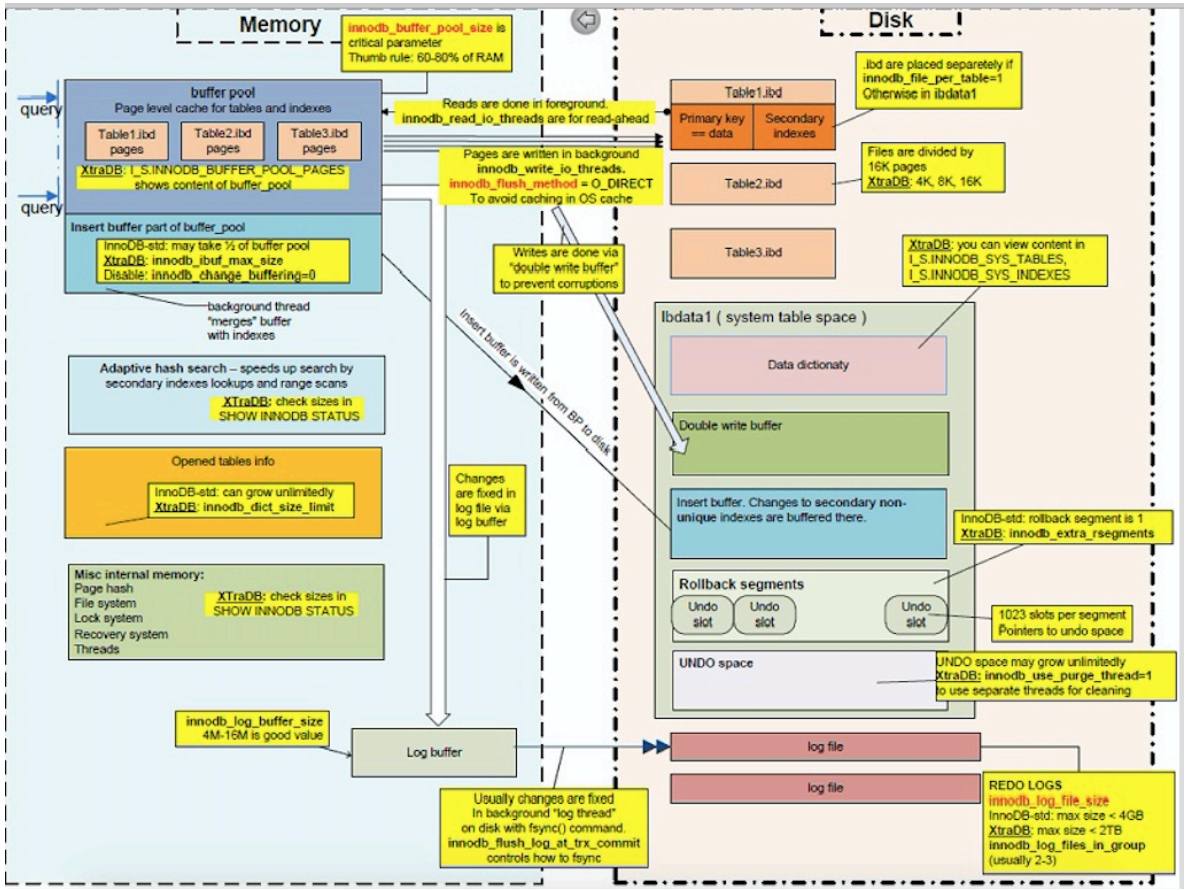
mysql -uroot -p1234 -e 'SHOW GLOBAL VARIABLES';

mysql -uroot -p1234 -e 'SHOW ENGINE INNODB STATUS’;

https://www.percona.com/about-us/mysql-white-paper/goal-driven-performance-optimization

Buffer	Cache
Container to hold data for a short period of time	A cache is a storage for speeding up certain operation
Buffer is normal speed storage	Cache is high speed storage area
Buffer is mostly used for I/O operation	Cache is used during R/W operation
Buffer is part of ram only	Cache is part of disk also
Buffer is made from dynamic ram	Cache is made from Static ram
Buffer's policy is first-in, first-out	Cache's policy is Least Recently Used

buffer , cache, pool



InnoDB Anatomy – Physical Structure