MySQL

Performance Management



MySQL Performance Cheat Sheet

Configuration Management



ClusterControl

Posted in:

Related resources:









MySQL is extensive and has lots of areas to optimize and tweak for the desired performance. Some changes can be performed dynamically, others require a server restart. It is pretty common to find a MySQL installation with a default configuration, although the latter may not be appropriate per se from your workload and setup.

Here are the key areas in MySQL which I have taken from different expert sources in the MySQL world, as well as our own experiences here at Severalnines. This blog would serve as your cheat sheet to tune performance and make your MySQL great again:-)

Let's take a look on these by outlining the key areas in MySQL.

System Variables

MySQL has lots of variables that you can consider to change. Some variables are dynamic which means they can be set using the SET statement. Others require a server restart, after they are set in the configuration file (e.g. /etc/my.cnf, etc/mysql/my.cnf). However, I'll go over the common things that are pretty common to tune to make the server optimized.

sort buffer size

This variable controls how large your filesort buffer is, which means that whenever a query needs to sort the rows, the value of this variable is used to limit the size that needs to be allocated. Take note that this variable is per-query that is processed (or per-connection) basis, which means that it would be a memory hungry when you set this higher and if you have multiple connections that



expected because data is retrieved randomly doing disk dives. This can cause performance degradation. However, it is best to fix your queries. Otherwise, if your application is designed to pull large queries and requires sorting, then it is efficient to use tools that handles query caching like Redis. By default, in MySQL 8.0, the current value set is 256 KiB. Set this accordingly only when you have queries that are heavily using or calling sorts.

read_buffer_size

MySQL documentation mentions that for each request that performs a sequential scan of a table, it allocates a read buffer. The read_buffer_size system variable determines the buffer size. It is also useful for MyISAM, but this variable affects all storage engines as well. For MEMORY tables, it is use to determine the memory block size.

Basically, each thread that does a sequential scan for a MyISAM table allocates a buffer of this size (in bytes) for each table it scans. It does applies for all storage engines (that includes InnoDB) as well, so it's helpful for queries that are sorting rows using ORDER BY and caching its indexes in a temporary file. If you do many sequential scans, bulk insert into partition tables, caching results of nested queries, then consider increasing its value. The value of this variable should be a multiple of 4KB. If it is set to a value that is not a multiple of 4KB, its value will be rounded down to the nearest multiple of 4KB. Take into account that setting this to a higher value will consume a large chunk of your server's memory. I suggest not to use this without proper benchmarking and monitoring of your environment.

read rnd buffer size

This variable deals with reading rows from a MyISAM table in sorted order following a key-sorting operation, the rows are read through this buffer to avoid disk seeks. The documentation says, when reading rows in an arbitrary sequence or from a MyISAM table in sorted order following a key-sorting operation, the rows are read through this buffer (and determined through this buffer size) to avoid disk seeks. Setting the variable to a large value can improve ORDER BY performance by quite a lot. However, this is a buffer allocated for each client, so you should not set the global variable to a large value. Instead, change the session variable only from within those clients that need to run large queries. However,

We use cookies on this site to enhance your user experience



By default, value is of 256K. The minimum size of the buffer that is used for plain index scans, range index scans, and joins that do not use indexes and thus perform full table scans. Also used by the BKA optimization (which is disabled by default). Increase its value to get faster full joins when adding indexes is not possible. Caveat though might be memory issues if you set this too high.

Remember that one join buffer is allocated for each full join between two tables. For a complex join between several tables for which indexes are not used, multiple join buffers might be necessary. Best left low globally and set high in sessions (by using SET SESSION syntax) that require large full joins. In 64-bit platforms, Windows truncates values above 4GB to 4GB-1 with a warning.

max_heap_table_size

This is the maximum size in bytes for user-created MEMORY tables are permitted to grow. This is helpful when your application is dealing with MEMORY storage engine tables. Setting the variable while the server is active has no effect on existing tables unless they are recreated or altered. The smaller of max_heap_table_size and tmp_table_size also limits internal in-memory tables. This variable is also in conjunction with tmp_table_size to limit the size of internal in-memory tables (this differs from the tables created explicitly as Engine=MEMORY as it only applies max_heap_table_size), whichever is smaller is applied between the two.

tmp_table_size

The largest size for temporary tables in-memory (not MEMORY tables) although if max_heap_table_size is smaller the lower limit will apply. If an in-memory temporary table exceeds the limit, MySQL automatically converts it to an on-disk temporary table. Increase the value of tmp_table_size (and max_heap_table_size if necessary) if you do many advanced GROUP BY queries and you have large available memory space. You can compare the number of internal on-disk temporary tables created to the total number of internal temporary tables created by comparing the values of the Created_tmp_disk_tables and Created_tmp_tables variables. In ClusterControl, you can monitor this via Dashboard -> Temporary Objects graph.

We use cookies on this site to enhance your user experience

By clicking the Accept button, you agree to us doing so. No, give me more info



value increases the number of file descriptors that mysqld requires, so you might as well consider checking your open_files_limit value or check how large is the SOFT and HARD limit set in your *nix operating system. You can monitor this whether you need to increase the table cache by checking the Opened_tables status variable. If the value of Opened_tables is large and you do not use FLUSH TABLES often (which just forces all tables to be closed and reopened), then you should increase the value of the table_open_cache variable. If you have a small value for table_open_cache, and a high number of tables are frequently accessed, this can affect the performance of your server. If you notice many entries in the MySQL processlistwith status "Opening tables" or "Closing tables", then it's time to adjust the value of this variable but take note of the caveat mentioned earlier. In ClusterControl, you can check this under Dashboards -> Table Open Cache Status or Dashboards -> Open Tables. You can check it here for more info.

table_open_cache_instances

Setting this variable would help improve scalability, and of course, performance which would reduce contention among sessions. The value you set here limits the number of open tables cache instances. The open tables cache can be partitioned into several smaller cache instances of size table_open_cache / table_open_cache_instances . A session needs to lock only one instance to access it for DML statements. This segments cache access among instances, permitting higher performance for operations that use the cache when there are many sessions accessing tables. (DDL statements still require a lock on the entire cache, but such statements are much less frequent than DML statements.) A value of 8 or 16 is recommended on systems that routinely use 16 or more cores.

table_definition_cache

Cache table definitions i.e. this is where the CREATE TABLE are cached to speed up opening of tables and only one entry per table. It would be reasonable to increase the value if you have large number of tables. The table definition cache takes less space and does not use file descriptors, unlike the normal table cache. Peter Zaitsev of Percona suggest if you can try the setting of the formula below,

1 $\|$ The number of user-defined tables + 10% unless 50K+ tables



So in case you have larger number of tables compared to the default, then it's reasonable you increase its value. Take into account that with InnoDB, this variable is used as a soft limit of the number of open table instances for the data dictionary cache. It will apply the LRU mechanism once it exceeds the current value of this variable. The limit helps address situations in which significant amounts of memory would be used to cache rarely used table instances until the next server restart. Hence, parent and child table instances with foreign key relationships are not placed on the LRU list and could impose a higher than the limit defined by table_definition_cache and are not subject to eviction in memory during LRU. Additionally, the table_definition_cache defines a soft limit for the number of InnoDB file-per-table tablespaces that can be open at one time, which is also controlled by innodb_open_files and in fact, the highest setting between these variables is used, if both are set. If neither variable is set, table_definition_cache, which has a higher default value, is used. If the number of open tablespace file handles exceeds the limit defined by table_definition_cache or innodb_open_files, the LRU mechanism searches the tablespace file LRU list for files that are fully flushed and are not currently being extended. This process is performed each time a new tablespace is opened. If there are no "inactive" tablespaces, no tablespace files are closed. So keep this in mind.



max_allowed_packet

This is the per-connection maximum size of an SQL query or row returned. The value was last increased in MySQL 5.6. However in MySQL 8.0 (at least on 8.0.3), the current default value is 64 MiB. You might consider adjusting this if you have large BLOB rows that need to be pulled out (or read), otherwise you can leave this default settings with 8.0 but in older versions, default is 4 MiB so you might take care of that in case you encounter ER_NET_PACKET_TOO_LARGE error. The largest possible packet that can be transmitted to or from a MySQL 8.0 server or client is 1GB.

skip_name_resolve



resolution and take advantage of improving your MySQL performance when this DNS lookup is disabled. Take into account that this variable is not dynamic, therefore a server restart is required if you set this in your MySQL config file. You may optionally start mysqld daemon, passing --skip-name-resolve option to enable this.

max connections

This is the number of permitted connections for your MySQL server. If you find out the error in MySQL 'Too many connections', you might consider setting it higher. By default, the value of 151 isn't enough especially on a production database, and considering that you have greater resources of the server (do not waste your server resources especially if it's a dedicated MySQL server). However, you must have enough file descriptors otherwise you will run out of them. In that case, consider adjusting your SOFT and HARD limit of your *nix operating systems and set a higher value of open_files_limit in MySQL (5000 is the default limit). Take into account that it is very frequent that the application does not close connections to the database correctly, and setting a high max_connections can result to some unresponsive or high load of your server. Using a connection pool at the application level can help resolve the issue here.

thread_cache_size

This is the cache to prevent excessive thread creation. When a client disconnects, the client's threads are put in the cache if there are fewer than thread_cache_size threads there. Requests for threads are satisfied by reusing threads taken from the cache if possible, and only when the cache is empty is a new thread created. This variable can be increased to improve performance if you have a lot of new connections. Normally, this does not provide a notable performance improvement if you have a good thread implementation. However, if your server sees hundreds of connections per second you should normally set thread_cache_size high enough so that most new connections use cached threads. By examining the difference between the Connections and Threads_created status variables, you can see how efficient the thread cache is. Using the formula stated in the documentation, 8 + (max_connections / 100) is good enough.

We use cookies on this site to enhance your user experience

By clicking the Accept button, you agree to us doing so. No, give me more info



MySQL 8.0 is that, the MySQL Team has stopped supporting this, as this variable can really cause performance issues. I have to agree on their blog that it is unlikely to improve predictability of performance. If you are engaged to use query caching, I suggest to use Redis or ProxySQL.

Storage Engine - InnoDB

InnoDB is an ACID-compliant storage engine with various features to offer along with foreign key support (Declarative Referential Integrity). This has a lot of things to say here but certain variables to consider for tuning:

innodb_buffer_pool_size

This variable acts like a key buffer of MyISAM but it has lots of things to offer. Since InnoDB relies heavily on the buffer pool, you would consider setting this value typically to 70%-80% of your server's memory. It is favorable also that you have a larger memory space than your data set, and setting a higher value for your buffer pool but not by too much. In ClusterControl, this can be monitored using our Dashboards -> InnoDB Metrics -> InnoDB Buffer Pool Pages graph. You may also monitor this with SHOW GLOBAL STATUS using the variables Innodb_buffer_pool_pages*.

innodb_buffer_pool_instances

For your concurrency workload, setting this variable can improve concurrency and reduce contention as different threads of read/write to cached pages. Minimum innodb_buffer_pool_instances should be lie between 1 (minimum) & 64 (maximum). Each page that is stored in or read from the buffer pool is assigned to one of the buffer pool instances randomly, using a hashing function. Each buffer pool manages its own free lists, flush lists, LRUs, and all other data structures connected to a buffer pool, and is protected by its own buffer pool mutex. Take note that this option takes effect only when innodb_buffer_pool_size >= 1GiB and its size is divided among the buffer pool instances.

innodb log file size

OK. I agree No. thanks



about: the recovery time after a crash. You need to balance recovery time in the rare event of a crash recovery versus maximizing throughput during peak operations. This limitation can translate to a 20x longer crash recovery process!

To elaborate it, a larger value would be good for InnoDB transaction logs and are crucial for good and stable write performance. The larger the value, the less checkpoint flush activity is required in the buffer pool, saving disk I/O. However, the recovery process is pretty slow once your database was abnormally shutdown (crash or killed, either OOM or accidental). Ideally, you can have 1-2GiB in production but of course you can adjust this. Benchmarking this changes can be a great advantage to see how it performs especially during after a crash.

innodb_log_buffer_size

To save disk I/O, InnoDB's writes the change data into It's log buffer and it uses the value of innodb_log_buffer_size having a default value of 8MiB. This is beneficial especially for large transactions as it does not need to write the log of changes to disk before transaction commit. If your write traffic is too high (inserts, deletes, updates), making the buffer larger saves disk I/O.

innodb_flush_log_at_trx_commit

When innodb_flush_log_at_trx_commit is set to 1 the log buffer is flushed on every transaction commit to the log file on disk and provides maximum data integrity but it also has performance impact. Setting it to 2 means log buffer is flushed to OS file cache on every transaction commit. The implication of 2 is optimal and improves performance if you can relax your ACID requirements, and can afford to lose transactions for the last second or two in case of OS crashes.

innodb_thread_concurrency

With improvements to the InnoDB engine, it is recommended to allow the engine to control the concurrency by keeping it to default value (which is zero). If you see concurrency issues, you can tune this variable. A recommended value is 2 times the number of CPUs plus the number of disks. It's dynamic variable means it can set without restarting MySQL server.

We use cookies on this site to enhance your user experience

By clicking the Accept button, you agree to us doing so. No, give me more info

several nines product DOWNI OAD PRICING PLANS **RESOURCES SUPPORT BLOG** LOGIN / SIGN UP SUBSCRIBE FOR NEWSLETTER Back to Home Page Filters Subscribe

read-heavy workload with mostly SELECT statements.

innodb file per table

innodb_file_per_table is ON by default from MySQL 5.6. This is usually recommended as it avoids having a huge shared tablespace and as it allows you to reclaim space when you drop or truncate a table. Separate tablespace also benefits for Xtrabackup partial backup scheme.

innodb stats on metadata

This attempts to keep the percentage of dirty pages under control, and before the Innodb plugin, this was really the only way to tune dirty buffer flushing. However, I have seen servers with 3% dirty buffers and they are hitting their max checkpoint age. The way this increases dirty buffer flushing also doesn't scale well on high io subsystems, it effectively just doubles the dirty buffer flushing per second when the % dirty pages exceeds this amount.

innodb_io_capacity

This setting, in spite of all our grand hopes that it would allow Innodb to make better use of our IO in all operations, simply controls the amount of dirty page flushing per second (and other background tasks like read-ahead). Make this bigger, you flush more per second. This does not adapt, it simply does that many iops every second if there are dirty buffers to flush. It will effectively eliminate any optimization of IO consolidation if you have a low enough write workload (that is, dirty pages get flushed almost immediately, we might be better off without a transaction log in this case). It also can quickly starve data reads and writes to the transaction log if you set this too high.

innodb_write_io_threads

Controls how many threads will have writes in progress to the disk. I'm not sure why this is still useful if you can use Linux native AIO. These can also be rendered useless by filesystems that don't allow parallel writing to the same file by more than one thread (particularly if you have relatively few tables and/or use the global tablespaces)

We use cookies on this site to enhance your user experience By clicking the Accept button, you agree to us doing so. No, give me more info

OK. I agree No. thanks

iiiioun_auaptive_iiusiiiig

Specifies whether to dynamically adjust the rate of flushing dirty pages in the InnoDB buffer pool based on the workload. Adjusting the flush rate dynamically is intended to avoid bursts of I/O activity. Typically, this is enabled by default. This variable, when enabled, tries to be smarter about flushing more aggressively based on the number of dirty pages and the rate of transaction log growth.

innodb_dedicated_server

This variable is new in MySQL 8.0 which is applied globally and requires a MySQL restart since it's not a dynamic variable. However, as documentation states that this variable is desired to be enabled only if your MySQL is running on a dedicated server. Otherwise, do not enable this on a shared host or shares system resources with other applications. When this is enabled, InnoDB will do an automatic configuration for the amount of memory detected for variables innodb_buffer_pool_size, innodb_log_file_size, innodb_flush_method. The downside only is that you cannot have the feasibility to apply your desired values on the detected variables mentioned.

MyISAM

key_buffer_size

InnoDB is the default storage engine now of MySQL, the default for key_buffer_size can probably be decreased unless you are using MyISAM productively as part of your application (but who uses MyISAM in production now?). I would suggest here to set perhaps 1% of RAM or 256 MiB at start if you have larger memory and dedicate the remaining memory for your OS cache and InnoDB buffer pool.

Other Provisions For Performance

slow_query_log

Of course, this variable does not help boost your MySQL server. However, this variable can help you out analyze slow performing queries. Value can be set to 0 or OFF to disable logging. Setting it to 1 or ON to enable this. The default value

We use cookies on this site to enhance your user experience

By clicking the Accept button, you agree to us doing so. No, give me more info

several nines product DOWNLOAD PRICING PLANS **RESOURCES SUPPORT** BLOG LOGIN / SIGN UP SUBSCRIBE FOR NEWSLETTER Back to Home Page Filters Subscribe

long_query_time

If a query takes longer than this many seconds, the server increments the Slow queries status variable. If the slow query log is enabled, the query is logged to the slow query log file. This value is measured in real time, not CPU time, so a query that is under the threshold on a lightly loaded system might be above the threshold on a heavily loaded one. The minimum and default values of long guery time are 0 and 10, respectively. Take note also that if variable min examined row limit is set > 0, it won't log queries even if it takes too long if the number of rows returned are less than the value set in min_examined_row_limit.

For more info on tuning your slow query logging, check the documentation here.

sync_binlog

This variable controls how often MySQL will sync binlogs to the disk. By default (>=5.7.7), this is set to 1 which means it will sync to disk before transactions are committed. However, this impose a negative impact on performance due to increased number of writes. But this is the safest setting if you want strictly ACID compliant along with your slaves. Alternatively, you can set this to 0 if you want to disable disk synchronization and just rely on the OS to flush the binary log to disk from time to time. Setting it higher than 1 means the binlog is sync to disk after N binary log commit groups have been collected, where N is > 1.

Dump/Restore Buffer Pool

It is pretty common thing that your production database needs to warm up from a cold start/restart. By dumping the current buffer pool before a restart, it would save the contents from the buffer pool and once it's up, it'll load the contents back into the buffer pool.. Thus, this avoids the need to warm up your database back to the cache. Take note that, this version was since introduced in 5.6 but Percona Server 5.5 has it already available, just in case you wonder. To enable this feature, set both variables innodb_buffer_pool_dump_at_shutdown = ON and innodb_buffer_pool_load_at_startup = ON.

Hardware

200 MHz.

For CPU, faster processors with multiple cores will be optimal for MySQL in most recent versions at least since 5.6. Intel's Xeon/Itanium processors can be expensive but tested for scalable and reliable computing platforms. Amazon has been shipping their EC2 instances running on ARM architecture. Though I personally haven't tried running or recall running MySQL on ARM architecture, there are benchmarks that had been made years ago. Modern CPU's can scale their frequencies up and down based on temperature, load, and OS power saving policies. However, there's a chance that your CPU settings in your Linux OS set to a different governor. You can check that out or set with "performance" governor by doing the following:

```
1 echo performance | sudo tee /sys/devices/system/cpu/cpu[0-9]*/cpufreq/scalir
```

For Memory, it is very important that your memory is large and can equate the size of your dataset. Ensure that you have swappiness = 1. You can check it out by checking sysctl or checking the file in procfs. This is achieved by doing the following:

```
1 $ sysctl -e vm.swappiness 2 vm.swappiness = 1
```

Or setting it to a value of 1 as follows

```
1 | $ sudo sysctl vm.swappiness=1 2 | vm.swappiness = 1
```

Another great thing to consider for your Memory management is considering turning off THP (Transparrent Huge Pages). In the past, I do recall we have some weird issues encountered with CPU utilization and thought it was due to disk I/O. It turned out, the problem was with kernel khugepaged thread which allocates memory dynamically during runtime. Not only this, during kernel goes for defragmentation, your memory will be quickly allocated as it passes it to THP. Standard HugePages memory is pre-allocated at startup, and does not change during runtime. You can verify and disable this by doing the following:

For Disk, it is important that you have a good throughput. Using RAID10 is the best

several nines product DOWNLOAD PRICING PLANS **RESOURCES SUPPORT** BLOG LOGIN / SIGN UP SUBSCRIBE FOR NEWSLETTER Back to Home Page Filters Subscribe

Operating system

Most production systems running on MySQL runs on Linux. It is because MySQL had been tested and benchmarked on Linux, and sounds that it's the de facto standard for a MySQL installation. However, of course, there's nothing stopping you from using it on Unix or Windows platform. It would be easier if your platform has been tested and there is a wide community to help, in case you experience some trouble. Most setups runs on RHEL/Centos/Fedora and Debian/Ubuntu systems. In AWS, Amazon has their Amazon Linux which I see as well being used in production by some.

Most important to consider with your setup is that your file system is using either XFS or Ext4. For sure, there are pros and cons between these two file systems but I won't go to the details here. Some say XFS outperform Ext4 but there are reports as well that Ext4 outperforms XFS. ZFS is also coming out of the picture as a good candidate for an alternative file system. Jervin Real (from Percona) has a great resource on this one, you can check this presentation during the ZFS conference.

External Links

https://developer.okta.com/blog/2015/05/22/tcmalloc

https://www.percona.com/blog/2012/07/05/impact-of-memory-allocators-onmysql-performance/

https://www.percona.com/live/18/sessions/benchmark-noise-reduction-how-toconfigure-your-machines-for-stable-results

https://zfs.datto.com/2018_slides/real.pdf

https://docs.oracle.com/en/database/oracle/oracle-database/12.2/ladbi/disablingtransparent-hugepages.html#GUID-02E9147D-D565-4AF8-B12A-8E6E9F74BEEA



Next blog >

We use cookies on this site to enhance your user experience By clicking the Accept button, you agree to us doing so. No, give me more info

OK. I agree No. thanks

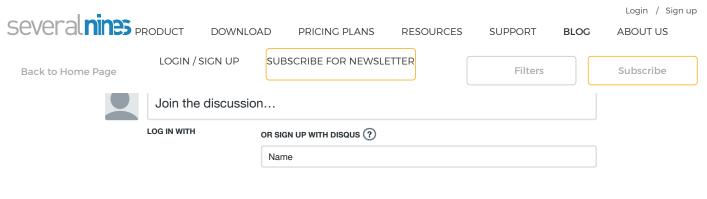
Login / Sign up



Paul Namuag has been able to work with various roles and is grateful to have a chance on different kinds of technologies for the past 18 years. He was a Graphic Artist and MS .Net Developer and switched to open source technologies since 2005 and was a Web Developer working with LAMP stack. After that, he was a Software Engineer/Game Engineer works with various companies developing for mobile or desktop and web applications. Later in 2013, he switched roles to a MySQL Support Engineer then Remote DBA at Percona which he was able to get a chance to understand how big data, highly scalable, highly available application works.

How to Deploy MariaDB Cluster 10.5 for High Availability	A Look at the Percona Toolkit
Dec 07, 2020 • by Paul Namuag	Nov 25, 2020 • by Paul Namuag

OK, I agree



Øystein Grøvlen • 2 years ago

Note that since MySQL 8.0.12, the sort buffer is allocated incrementally (https://dev.mysql.com/workl.... Hence, increasing the sort_buffer_size should not have any significant impact on queries that only require a small sort buffer.

^ | ✓ • Reply • Share •

Paul Namuag → Øystein Grøvlen • 2 years ago

Thank you for pointing that Øystein. Would it mean to have sort_buffer_size be deprecated and instead MySQL will automatically resized it if needed in the future? Although if not, it could make more sense though to still avoid such larger value if most queries running or executed are design to run heavy sorting queries.

∧ | ∨ • Reply • Share •

Øystein Grøvlen → Paul Namuag • 2 years ago

I think one will still want to be a able to set a maximum size for the sort buffer. Otherwise, queries that sort a lot of data may allocate more than the available memory.

8 ^ | V · Reply · Share ›

The only management system you'll ever need to take control of your open source database infrastructure.

Learn more







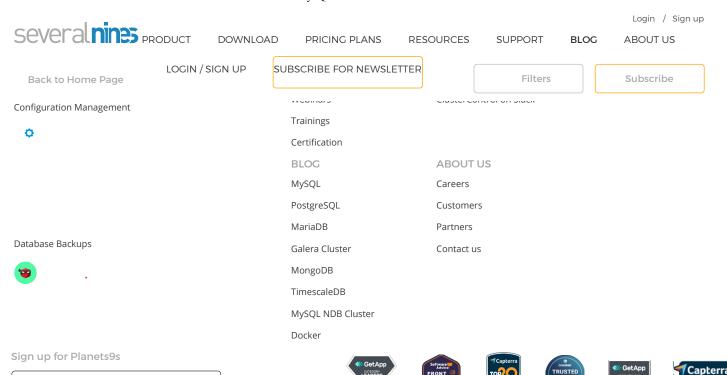












Your e-mail

Subscribe













© Copyright 2014-2020 Severalnines AB. All rights reserved.

Website Policies | Sitemap