

[MySQL] Conexiones a bases de datos desde VPN

Cambios en las conexiones desde VPN hacia bases de datos MySQL

Actualmente desde nuestras computadoras personales nos estamos conectando mediante VPN con nuestro usuario nominal de forma directa a las bases de datos MySQL.

Estas conexiones directas serán bloqueadas y sólo se permitirá conectarse a las bases de datos MySQL desde VPN a través de un proxy implementado por el equipo de Seguridad Informática (utilizando la herramienta open source ProxySQL).

No es necesario instalar software alguno, simplemente apuntar las conexiones al nuevo host como se detalla en la siguiente sección. Podés seguir utilizando cualquier cliente MySQL de tu preferencia ya que siguen siendo compatibles (por ejemplo: MySQL Workbench, SequelPro, DBeaver, Datagrip, etc.)

A partir del día 12/05/2020 el proxy queda disponible para ser utilizado por cualquier usuario, y el día 22/06/2020 se realizará el bloqueo de las conexiones directas desde VPN (*notar que se repite varias veces la palabra VPN para dejar en claro que no afecta a ninguna aplicación corriendo en cloud*).

Antes de contactarnos, por favor lean el documento y los errores comunes al intentar conectarse, qué *casi siempre* son 2:

- Error de usuario, contraseña o esquema introducidos.
- Error de conectividad (abajo se detalla qué deben hacer en este caso, si contactan a datasec no podremos ayudarlos porque no gestionamos los firewalls de red).

Conexión a través del proxy

Para conectarse a ProxySQL, los datos de conexión son los siguientes:

- Host:
 - proxysql.master.meliseginf.com (para conectarse a las productivas)
 - proxysql.slave.meliseginf.com (para conectarse a las réplicas)
 - proxysql.slave.ohio.meliseginf.com (para conectarse a las réplicas de Ohio)
 - proxysql.clone.meliseginf.com (para conectarse a los clone si crearon alguno)
- Puerto: 6612
- Database: el nombre del esquema al que quieran conectarse (el esquema, *NO el nombre del cluster/base de datos generado en Fury*).

El usuario y contraseña que deben ingresar es el de su usuario nominal, este proxy no permite conexiones con ningún usuario aplicativo (como por ejemplo los _RPROD, _WPROD y _ADMIN que se crean con cada esquema).

Para actualizar una conexión existente, simplemente debemos apuntar al nuevo host. Por ejemplo:











Conexión a través de ProxySQL

¿Qué sucede si no completo un esquema al conectarme?

Si en la conexión no especificas el esquema, por defecto ProxySQL te llevará a la base de datos que tengas seleccionada por defecto en

https://dbaccess.meliseginf.com/#!/mysql/my_accesses en la columna “Conexión default”:

MySQL - Gestión de mis accesos									
CAMBIAR CONTRASEÑAS									
Acciones	Esquema ▾	Base de datos ▾	Owner principal ▾	Owners secundarios ▴	Lectura	Escritura	Admin	Conexión default	
	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>		
	beeClas	beeclasifica00	rmarzovilla	ekim	SI 	NO 	NO 	<input type="checkbox"/>	
	findings	beefindings00	raguirre	ekim	SI 	NO 	NO 	<input checked="" type="checkbox"/>	

MySQL - Gestión de accesos en DBAccess

Desde ahí mismo podés seleccionar otro esquema por defecto.

¿Con qué contraseña me conecto?

La contraseña es única para conectarse a todas las bases de datos MySQL, y la podés cambiar desde el botón superior derecho “CAMBIAR CONTRASEÑAS” en

https://dbaccess.meliseginf.com/#!/mysql/my_accesses.

Próximamente ese cambio de contraseña podrá estar integrado en Mordor para que puedas utilizar tu misma contraseña de red para acceder a las bases de datos MySQL.

¿Cómo sabe ProxySQL a qué base de datos conectarme con solo especificar el esquema?

En ProxySQL hay reglas cargadas para que, según el esquema especificado en la conexión, te derive automáticamente al servidor de base de datos correcto. Es por ello que se hace posible que ProxySQL escuche conexiones en 1 único puerto para todas las bases de datos.

Todos los servidores, esquemas y usuarios nominales son cargados y actualizados de forma automática desde DBAccess.

Como el ruteo se basa en el nombre del esquema, en una conexión existente con la siguiente sentencia podés cambiar hacia otras bases de datos a las que tengas acceso:

```
USE esquema;
```

Preguntas y errores frecuentes

¿Debo configurar este proxy en las aplicaciones que están corriendo en Fury?

NO, como se especificó anteriormente, este proxy debe utilizarse sólo para acceder desde tu máquina personal conectado a la VPN y con tu usuario nominal. Este cambio no tiene ningún impacto en las aplicaciones. Además no podrás conectarte al proxy con ningún usuario aplicativo.

¿Qué bases de datos están incluidas en este esquema?

Todas las bases de datos MySQL gestionadas por DBA creadas a través del servicio DBaaS, independientemente de si están catalogadas como de test o no.

¿Esto incluye a todos los empleados?

Cualquier empleado que quiera conectarse a una base de datos MySQL DBaaS deberá hacerlo a través del proxy, porque las reglas de los firewalls de red no permitirán la conexión directa.

¿Las queries que ejecuto sobre una base de datos son monitoreadas?

Todas las queries ejecutadas por cada usuario en cualquier base de datos quedan registradas, tanto con las conexiones a través del proxy como con las directas. Por lo cual es importante que sepas que tu usuario nominal es personal e intransferible (no debes prestarlo a otras personas).

Si estoy utilizando Tableau Desktop, ¿cómo configuro la conexión?



The image shows a 'MySQL' connection dialog box with a close button (X) in the top right corner. The fields are as follows:

- Servidor:** master.meliseginf.com
- Puerto:** 6612
- Base de datos:** dbaccess
- Escriba la información para iniciar sesión en la base de datos:**
 - Nombre de usuario:** rmarzovilla
 - Contraseña:** (empty field)
- ☐ **Requiere SSL**
- SQL inicial...** (link)
- Iniciar sesión** (button)

Conexión de Tableau Desktop a MySQL

De acuerdo a la versión de Tableau Desktop que tengas podrás especificar el nombre del esquema en el campo “Base de datos”.

Si no tenés ese campo para completar, podés actualizar la versión de la aplicación ó haciendo click en el botón inferior izquierdo “SQL inicial...” añadir la siguiente sentencia:

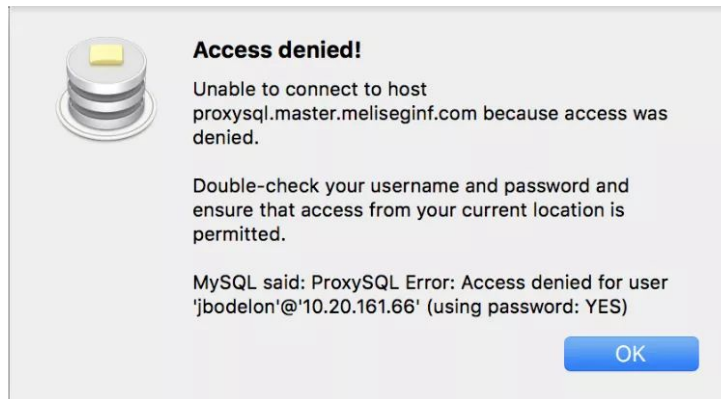
```
USE esquema;
```

Al publicar el tablero en Tableau Server, si usan el proxy + usuario nominal igualmente va a funcionar, pero lo *recomendable* es configurar la conexión directa a la réplica de la base de datos (cluster.slave.mlaws.com) con el usuario aplicativo de solo lectura (*_RPROD) así si cambiás la contraseña o tu usuario se elimina no se ve afectado el tablero, y además usando la slave no generarás tráfico en la instancia master. Esto mismo aplica para cualquier proceso o aplicación en cloud, no deberían estar corriendo con usuarios nominales.

El usuario aplicativo *_RPROD le tuvo que haber llegado a través de un email automático de Fury a quién creó el esquema. En caso de no contar con él, el owner principal del esquema puede obtenerlo desde DBAccess en https://dbaccess.meliseginf.com/#!/mysql/my_schemas.

Al conectarme me arroja el error: “Access denied for user...”

Si al intentar conectarte te arroja el error Access Denied:

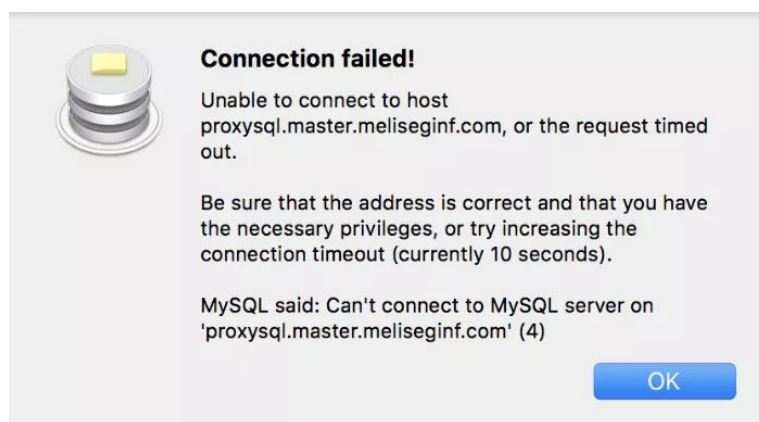


Error Access Denied al conectarse a MySQL

Verificá que el usuario y contraseña que estás ingresando sea el correcto, sino probá cambiando tu contraseña.

Al intentar conectarme me arroja error de conectividad (timeout)

Si al intentar conectarte te arroja el error “Can’t connect to MySQL server on...”:



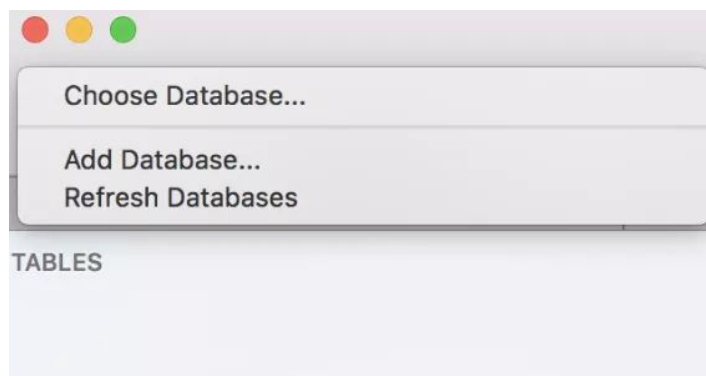
Error de conectividad a MySQL

Verificá estar conectado a la VPN, cerrar cualquier máquina virtual abierta y pasar los chequeos del estado de salud de tu equipo en [selfservice](#).

Si persiste el problema de conectividad, cargar ticket en Shield en la categoría "VPN (GlobalProtect) - Perdida de accesos" ó "VPN (GlobalProtect) - Nuevos Accesos" y el equipo de Internal Systems de infraops analizará el caso.

Me conecté al proxy pero no me carga ningún esquema, ¿qué puede estar pasando?

Si al conectarte al proxy no te carga ningún esquema y tablas:



Error de acceso a un esquema no válido o sin permisos

Significa que ingresaste un esquema inválido o un esquema sobre el que no tenés permisos.

Revisá haber completado correctamente el nombre del esquema (no el nombre del cluster) y que tengas permisos sobre ese esquema en https://dbaccess.meliseginf.com/#!/mysql/my_accesses (sólo podés acceder a aquellas bases donde tengas acceso).

Ante cualquier otra consulta o sugerencia, nos pueden contactar a:
datasec@mercadolibre.com

Buenas practicas para alta performance en MySQL

Hola a todos!

En el siguiente articulo enumeraremos una serie de buenas practicas para optimizar las consultas que impactan contra las bases de datos.

Diseño de tablas

- Crea siempre tus tablas con Primary Key **autoincrement bigint unsigned**
- Intenta evitar el uso de Foreign Keys debido a que penalizan la escritura.
- Usa InnoDB como engine a la hora de crear la tabla (Default).
- Recomendamos el uso de particiones cuando las tablas son muy grandes. (Ver documentación)
- Si querés estimar cuanto demora en crearse un indice o un alter podes usar un clon (Ver documentación)
- Para columnas tipo Varchar intentemos siempre acotar el ancho al definirlas, evitemos usar Varchar(255) para reducir impacto en memoria
- Tener en cuenta los tipos de datos que usamos en columnas que luego estarán involucradas en predicados tanto en JOINS como en WHERE. Si comparamos columnas con tipos distintos obligamos al motor a realizar un type cast implícito por cada comparación.

<https://www.oreilly.com/library/view/high-performance-mysql/9781449332471/ch04.html>

Distribuir el tráfico

- Podes enviar las lecturas a los slaves y así dejar el master con sus recursos para lo realmente importante. Ejemplo: cambiar el endpoint `clustername.master.mlaws.com` por `clustername.slave.mlaws.com`.

Pool de conexiones

- Siempre asegurarnos de estar utilizando un pool de conexiones. Si nuestra app tiene que conectarse a la DB para cada query a ejecutar, vamos a sufrir un performance penalty altísimo y según el caso podríamos alcanzar el máximo permitido de conexiones con facilidad.
- Configurar el pool de conexiones de manera tal que la operación normal de nuestro scope no sea con porcentajes altos de utilización del número conexiones en el server. Esta suele ser una causa de fallas de deploys Blue/Green

Acá va un artículo del creador de Hikari analizando teóricamente y experimentalmente la performance de distintos sizes de pool de conexiones.

<https://github.com/brettwooldridge/HikariCP/wiki/About-Pool-Sizing>

Aplica para cualquier pool en general, no solo Hikari!

Performance tools

- Podes usar la herramienta **EXPLAIN** para ver el plan de ejecución de la consulta.

Ejemplo de query donde se lee toda la tabla (**FULL SCAN**): Type=ALL y Key=NULL

```
mysql> explain select * from movimientos;
```

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	movimientos	NULL	ALL	NULL	NULL	NULL	NULL	998362	100.00	NULL

1 row in set, 1 warning (0.01 sec)

Ejemplo de query donde se accede a la tabla por un índice:

```
mysql> explain select id, importe from movimientos where idcliente=1;
```

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	movimientos	NULL	ref	idx_movimientos_cli	idx_movimientos_cli	4	const	212106	100.00	NULL

1 row in set, 1 warning (0.00 sec)

Ejemplo de query con join de dos tablas. El optimizador ignora la PK de la tabla clientes porque tiene muy pocos registros y accede a la segunda por índice:

```
mysql> explain select c.nombre, c.apellidos, count(*) from movimientos m inner join clientes c on (c.id=m.idcliente) where c.apellidos='Perez' group by c.nombre, c.apellidos;
```

id	select_type	table	partitions	type	possible_keys	key	key_len	ref	rows	filtered	Extra
1	SIMPLE	c	NULL	ALL	PRIMARY	NULL	NULL	NULL	10	10.00	Using where; Using temporary
1	SIMPLE	m	NULL	ref	idx_movimientos_cli	idx_movimientos_cli	4	demodba.c.id	110929	100.00	Using index

2 rows in set, 1 warning (0.00 sec)

Possible keys: Nos muestra los posibles índices que se pueden utilizar en la consulta

Key: Nos muestra el índice que está utilizando para la consulta. Si su valor es NULL, quiere decir que nos está faltando un índice apropiado para la consulta.

Rows: Muestra la cantidad de filas que el motor va a utilizar para realizar la consulta.

Para más información sobre la herramienta EXPLAIN podes consultar la documentación oficial aca.

- Trata de evitar los full scan de tablas. Si estas son chicas pueden ser aceptables pero igualmente si la carga es alta esto puede generar contención a la hora de acceder a dicha tabla.
- Si el acceso al índice es del tipo RANGE SCAN o FULL SCAN puede implicar una gran cantidad de lecturas al índice generando un bajo rendimiento.
- Podés revisar si tenés Slow Queries desde Fury. (Ver documentación)
- También puedes ver el estado de tus sesiones con el Processlist desde DBAccess. (Ver documentación)

Buenas prácticas a la hora de acceder una tabla

Índices:

- Es muy importante tener sólo los índices necesarios, ya que su uso penaliza la escritura e incrementa el espacio utilizado por la base de datos y el consumo de cpu.
- Si creamos índices compuestos, el orden es sumamente importante. Deben ordenarse desde las columnas más restrictivas a las menos, es decir de la que más filtra a la que menos filtra (ej: site_id, fecha es mejor que fecha, site_id). Igualmente tené en cuenta que si querés recorrer la mayor parte de la tabla, el optimizador va a realizar un full access.
- Trata de evitar SELECT * FROM. Tratar de que se usen los campos del índice o solo los campos necesarios en tu consulta.
- Intenta evitar funciones u operadores sobre una columna en el WHERE ya que provoca que el optimizador ignore los índices de dicha columna. Ejemplo:

```
... WHERE order.date_created BETWEEN '2020-04-02' and '2020-04-03'
```

performa mejor que:

```
... WHERE DATE(order.date_created) = '2020-04-02'
```

- Trata de evitar índices de función ya que su costo es alto.
- Te sugerimos que evites el uso el operador desigual (<> o !=) ya que puede provocar que se ignoren los índices por ese campo.

Consideraciones para trabajar con grandes volúmenes de datos

- Tené en cuenta el volumen de datos de la consulta que queremos realizar. Ejecutar la siguiente consulta para verlo:

```
SELECT table_schema AS "Schema", table_name AS "Table",
round(((data_length + index_length) / 1024 / 1024 / 1024), 2) "Size in GB"
FROM information_schema.TABLES
WHERE table_schema = 'SCHEMA_NAME'
ORDER BY (data_length + index_length) DESC;
```

- Evitar consultas sin where.
- Evitar **ORDER BY** sobre consultas muy grandes.
- Evitar el uso de **LIKE** en el **WHERE** de las consultas.
- Evitar el uso de **DISTINCT**.
- Para realizar updates masivos te recomendamos la utilización de cursores (Ver ejemplo)
- Si tenés procesos que usan tablas "temporales" o "auxiliares", que son habitualmente llenadas y vaciadas (ej. diariamente) y que no son utilizadas por la aplicación, ponete en contacto con nosotros para skippearlas en los slaves.
- Si vas a ejecutar varios alters sobre una tabla, te recomendamos que lo hagas en un mismo comando para corra más rápido. Ej:

```
ALTER TABLE ejemplo ADD column_name column_type, ADD  
INDEX(column_name);
```

Esperamos que les sea de utilidad

Saludos!

DBA Team