

Universidad Francisco Marroquin
Facultad de Ciencias Economicas
Computer Science
Datos I
Profesor: Luis Fernando Alonso
Sophia Gamarro: 20170356

Herramientas de rendimiento y SLAs

MySQL

MySQL: MySQL es un sistema de administración de bases de datos relacionales de código abierto basado en Structured Query Language (SQL). Los equipos de desarrollo utilizan MySQL en una amplia variedad de casos de uso, más comúnmente en aplicaciones de almacenamiento de datos, comercio electrónico y registro. Sin embargo, su implementación más popular es en bases de datos web, ya que es altamente flexible, confiable y escalable. Se puede utilizar para almacenar todo, desde registros únicos de información hasta inventarios completos de productos.

Medir el rendimiento de la base de datos MySQL

En general, las métricas de rendimiento de la base de datos se dividen en dos tipos principales: métricas de carga de trabajo y métricas de recursos. Las métricas de carga de trabajo miden el rendimiento general de una base de datos midiendo la cantidad de trabajo o salida que puede producir en un período determinado. Mientras tanto, las métricas de recursos miden cuánto hardware, software y recursos de red consume una base de datos para funcionar.

Para la mayoría de los casos de uso, las categorías más importantes son:

- Rendimiento: métricas de carga de trabajo que miden cuánto trabajo realiza una base de datos en un intervalo de tiempo determinado (por ejemplo, número de consultas, transacciones, lecturas, escrituras, etc.)

- Latencia (tiempo de ejecución): métricas de carga de trabajo que miden el tiempo que tarda una base de datos en realizar una unidad de trabajo determinada (p. Ej., Tiempos de ejecución de consultas)
- Conexiones (simultaneidad): métricas de recursos que le indican la cantidad de consultas que se ejecutan simultáneamente y que no se pueden completar a tiempo (por ejemplo, cantidad de subprocesos conectados, subprocesos en ejecución, conexiones abortadas, etc.)
- Búfer (utilización): métricas de recursos que miden el uso del búfer (caché) (p. Ej., Utilización del grupo de búfer)

Ubicación de las métricas de rendimiento de MySQL

Las métricas de supervisión del rendimiento de MySQL se pueden consultar desde tres ubicaciones diferentes: variables de estado del servidor, esquema del sistema y esquema de rendimiento.

Variables de estado del servidor: Son “contadores” internos que MySQL mantiene para brindar información sobre su funcionamiento. Estas variables incluyen Preguntas, Slow_queries, Threads_running, Com_select y otras.

Esquema de rendimiento: esta es una función de MySQL para monitorear los eventos de MySQL Server y la ejecución de consultas. Las métricas de rendimiento se obtienen consultando una o más tablas en la base de datos performance_schema, como la tabla events_statements_summary_by_digest.

Esquema del sistema: el esquema del sistema de MySQL mejora la usabilidad del esquema de rendimiento (que, con docenas de tablas, a veces puede ser tedioso para trabajar). El esquema incluye vistas, funciones y procedimientos que hacen que sea más conveniente analizar los datos recopilados por el esquema de rendimiento.

Métrica	Nombre Mostrado de Métrica	Unidad	Descripción
CurrentConnections	Conexiones actuales	recuento	Número total de conexiones a MySQL Database.
ActiveConnections	Conexiones Activas	recuento	Número de conexiones activas a MySQL Database.
Sentencias	Recuento de extractos	recuento	Número de sentencias ejecutadas en MySQL Database.
StatementLatency	Latencia de extracto	ms	Latencia de extracto para todas las sentencias ejecutadas.
CPUUtilization	Uso de CPU	porcentaje	Uso de CPU para el host de MySQL Database.
MemoryUtilization	Uso de Memoria	porcentaje	Uso de Memoria para el Host de MySQL Database.
DbVolumeReadOperations	Operaciones de Lectura de Disco	recuento	Número total de operaciones de lectura para los volúmenes de base de datos.
DbVolumeWriteOperations	Operaciones de escritura de disco	recuento	Número total de operaciones de escritura para los volúmenes de base de datos.

¿Por qué es útil supervisar el rendimiento de la base de datos MySQL?

Supongamos que mi base de clientes está creciendo o me encuentro experimentando problemas de rendimiento frecuentes con la implementación de MySQL, una herramienta de monitoreo de base de datos es necesaria.

Las herramientas de monitoreo de MySQL logran esto a través de métricas del sistema de bajo nivel que ofrecen una descripción general de su implementación. Una buena herramienta de monitoreo también enviará alertas cuando las métricas se desvíen de un estado estable para que pueda intervenir rápidamente y evitar posibles interrupciones u otros problemas que afecten al cliente.

Capacidades de monitoreo de MySQL

La supervisión de su base de datos MySQL permite crear una imagen del rendimiento del sistema. Las siguientes categorías de estadísticas son esenciales para determinar si MySQL está en buen estado:

- Errores
- Latencia
- Saturación de recursos
- Estado del disco
- Retraso de replicación
- Conexiones máximas
- Fuentes de latencia y consultas lentas

Algunas herramientas populares y muy útiles que puede utilizar my MySQL para dicho monitoreo son las siguientes:

1. MySQL Enterprise Monitor
 - a. Esta herramienta de monitoreo de MySQL es capaz de detectar y mapear configuraciones de clústeres y monitoreará el estado de los servicios y servidores de soporte, así como las propias instancias de la base de datos. También rastrea la topología de replicación y administra las funciones de respaldo y restauración. Observa las conexiones que utilizan las aplicaciones para acceder a la base de datos MySQL e informa sobre los tiempos de respuesta y el uso de recursos.

2. Datadog MySQL

- a. Datadog es una solución de monitorización basada en la nube que está disponible en varios módulos, siendo la monitorización MySQL parte de su servicio APM. Su agente de integración de MySQL sondea su esquema de rendimiento de MySQL y emite las métricas, que luego se generan y se muestran en los paneles de Datadog.

3. Percona Monitoring and Management Tool (PMM):

PMM es un conjunto de herramientas de monitoreo de código abierto construido sobre el ecosistema de Prometheus. Proporciona paneles de control listos para usar para ayudarlo a comprender el rendimiento de su sistema MySQL. PMM aprovecha Prometheus a través de la herramienta MySQLD para extraer métricas de MySQL. Mysqld_exporter requiere acceso para consultar sus instancias de MySQL.

4. Prometheus and MySQLD Exporter

Prometheus es una herramienta de supervisión de código abierto inspirada en la solución de supervisión interna de Google. Es extremadamente popular, con más de 33,000 estrellas en GitHub. Este monitor de base de datos requiere ejecutar el servicio MySQLD Exporter, que consulta a MySQL para obtener estadísticas de rendimiento. También requiere ejecutar el servicio Prometheus, así como una herramienta de visualización independiente como Grafana.

5. Sematext MySQL Monitoring

- a. El descubrimiento automático de servicios de Sematext le permite comenzar a monitorear automáticamente sus instancias de MySQL directamente a través de la interfaz de usuario.

PostgreSQL

PostgreSQL: es una base de datos relacional de código abierto de clase empresarial avanzada que admite consultas tanto SQL (relacional) como JSON (no relacional). Es un sistema de administración de bases de datos altamente estable, respaldado por más de 20 años de desarrollo comunitario que ha contribuido a sus altos niveles de resiliencia, integridad y corrección. PostgreSQL se utiliza como almacén de datos principal o datawarehouse para muchas aplicaciones web, móviles, geoespaciales y de análisis.

PostgreSQL recopila automáticamente una cantidad sustancial de estadísticas sobre su actividad, pero la investigación se centrará en solo algunas categorías de métricas que pueden ayudar a obtener información sobre el estado y el rendimiento de sus servidores de base de datos:

Todas las métricas mencionadas son accesibles a través del recopilador de estadísticas de PostgreSQL y otras fuentes nativas.

- Tiempo de respuesta

El tiempo de respuesta es un indicador importante del rendimiento y la eficiencia general de un servidor de base de datos. Es el tiempo que tarda el servidor en responder a un determinado comando / consulta desde el momento en que se emitió.

- Estadísticas de conexión

El servidor puede tener varias conexiones activas ejecutándose simultáneamente en la base de datos.

La herramienta de monitoreo de desempeño PostgreSQL de Applications Manager ayuda a detectar un desempeño lento al monitorear el número de conexiones activas. Si la cantidad de conexiones es alta, es posible que deba analizar el estado de estas sesiones de usuario y finalizar las sesiones inactivas que ralentizan el servidor.

- Estadísticas de bloqueo: PostgreSQL proporciona varios modos de bloqueo para controlar el acceso simultáneo a los datos de las tablas. Este mecanismo garantiza la coherencia de los datos en la base de datos.
- Estadísticas de búfer: Las operaciones de la base de datos suelen ser más rápidas cuando los datos se leen de la memoria. PostgreSQL ofrece diferentes tipos de búfer, como búfer compartidos, búfer WAL y búfer CLOG. Estos búferes tienen capacidades y funciones variadas. Por ejemplo, los búferes compartidos almacenan datos para un acceso rápido, los búferes WAL almacenan metadatos sobre cambios en los datos reales para operaciones de recuperación de bases de datos y los búferes CLOG contienen páginas de registro de confirmación.

- Detalles de la sesión y la consulta: Monitorear las sesiones y las consultas puede brindarle una comprensión de las operaciones que se están ejecutando en la base de datos.
- Detalles de la replicación: La replicación es un proceso en el que los datos se copian (sincronizan) desde una base de datos en un servidor (maestro) a una base de datos en otro servidor (esclavos). PostgreSQL ofrece un servicio de replicación de transmisión interna que crea un entorno de alta disponibilidad, equilibra la carga de consultas de solo lectura en varios nodos, crea una base de datos de solo lectura para ejecutar consultas analíticas y muchas otras funciones fundamentales. La supervisión de las réplicas es un aspecto fundamental de la supervisión de PostgreSQL, ya que las réplicas a veces pueden desincronizarse.

Nombre	Tipo	Unidad	Tipo de agregación	Descripción
Porcentaje de CPU	Métrica	Porcentaje	Promedio.	Porcentaje de CPU en uso.
Porcentaje de memoria	Métrica	Porcentaje	Promedio.	Porcentaje de memoria en uso.
Porcentaje de E/S	Métrica	Porcentaje	Promedio.	Porcentaje de E/S en uso.
Porcentaje de almacenamiento	Métrica	Porcentaje	Promedio.	Porcentaje de almacenamiento en uso del máximo del servidor.
Almacenamiento en uso	Métrica	Bytes	Promedio.	Cantidad de almacenamiento en uso. El almacenamiento utilizado por el servicio incluye los archivos de base de datos, los registros de transacciones y los registros del servidor.
Límite de almacenamiento	Métrica	Bytes	Promedio.	Almacenamiento máximo para el servidor.
Porcentaje de almacenamiento de registro del servidor	Métrica	Porcentaje	Promedio.	Porcentaje de almacenamiento de registro del servidor en uso del máximo del almacenamiento de registro del servidor.
Almacenamiento de registro del servidor usado	Métrica	Bytes	Promedio.	Cantidad de almacenamiento de registro del servidor en uso.
Límite de almacenamiento de registro del servidor	Métrica	Bytes	Promedio.	Almacenamiento máximo de registro del servidor para el servidor.
Conexiones activas	Métrica	Count	Promedio.	Número de conexiones activas con el servidor.
Conexiones erróneas	Métrica	Count	Total	Número de conexiones con errores con el servidor.
Almacenamiento de copia de seguridad usado	Métrica	Bytes	Promedio.	Cantidad de almacenamiento de copia de seguridad que se utiliza.
Salida de red	Métrica	Bytes	Total	Salida de red en conexiones activas.
Entrada de red	Métrica	Bytes	Total	Entrada de red en conexiones activas.
Retraso de réplica	Métrica	Segundos	Máxima	Número de segundos que el servidor de réplica va con retraso respecto del servidor primario.
Retraso máximo entre réplicas	Métrica	Bytes	Máxima	Retraso en bytes del servidor de réplica más atrasado.

Algunas herramientas populares y muy útiles que puede utilizar PostgreSQL para dicho monitoreo son las siguientes:

- Monitor de rendimiento de la base de datos SolarWinds DPM, en segundo lugar cercano a Datadog, proporciona un seguimiento detallado y una automatización proactiva.

- Monitor de red de PRTG: viene con un sensor PostgreSQL específico y es ideal para quienes necesitan un control adicional.
- Datadog APM: el mejor monitor PostgreSQL completo con implementación rápida y conocimientos listos para usar.
- AppOptics APM: incluye utilidades de optimización y monitoreo de permanencia.
- Administrador de aplicaciones ManageEngine: puede administrar y monitorear múltiples RDBMS y servidores.
- SQL Sentry: dedicado a la gestión de entornos SQL. Cuenta con herramientas de configuración e implementación.
- ClusterControl: disponible como herramienta gratuita y de pago.

SLA para servicios de base de datos administrados en la nube

1. Oracle Autonomous Database: Solo Oracle ofrece SLA de extremo a extremo que cubren el rendimiento, la disponibilidad y la capacidad de gestión.
 - Disponibilidad / plano de datos: los servicios están en funcionamiento con tiempo de actividad y conectividad comprometidos
 - Plano de control / capacidad de administración: capacidad para administrar, monitorear y modificar los recursos de infraestructura de ORACLE CLOUD.
 - Rendimiento: los servicios se desempeñan consistentemente como se esperaba

Con respecto a la base de datos de Oracle Cloud, el SLA de Oracle = (100% - lo que ellos definen como "el promedio de las Tasas de Error de cada período de una hora en un mes determinado del Período de Servicios").

2. AWS RDS El acuerdo de nivel de servicio de Amazon RDS cubre implementaciones de varias zonas de disponibilidad de instancias de servicio de base de datos relacional para los siguientes motores de base de datos alojados en RDS:
 - MySQL
 - MariaDB
 - Oracle
 - PostgreSQL
 - servidor SQL

SLA de tiempo de actividad de Amazon RDS

- Para todas las instancias de RDS alojadas en varias zonas de disponibilidad (con el parámetro "Multi AZ" establecido en "Verdadero"), Amazon garantiza un tiempo de actividad del 99,5% en cualquier ciclo de facturación mensual. Esto permite hasta 3,65 horas de tiempo de inactividad por mes.
- Créditos de servicio de Amazon RDS
- Al igual que con EC2 y todos los servicios de AWS, los créditos de servicio deben solicitarse para utilizar el procedimiento descrito anteriormente. Los créditos de servicio para RDS son los siguientes:

Monthly Uptime Percentage	Service Credit Percentage	Downtime Per Month
Less than 99.95% but equal to or greater than 99.0%	10%	3.65 hours to 7.31 hours
Less than 99.0% but equal to or greater than 95.0%	25%	7.31 hours to 36.53 hours
Less than 95.0%	100%	More than 36.53 hours

3. Azure SQL Database: Azure SQL Database es una base de datos relacional totalmente administrada con alta disponibilidad regional integrada y replicación geográfica llave en mano para cualquier región de Azure. Incluye inteligencia para admitir funciones de conducción autónoma, como ajuste de rendimiento, monitoreo de amenazas y evaluaciones de vulnerabilidad, y proporciona parches y actualizaciones completamente automatizados de la base de código.

- Los niveles Business Critical o Premium de Azure SQL Database configurados como implementaciones con redundancia de zona tienen una garantía de disponibilidad de al menos el 99,995%.
- Los niveles Business Critical o Premium de Azure SQL Database no configurados para implementaciones con redundancia de zona, niveles de uso general, estándar o básico, o nivel de hiperescala con dos o más réplicas tienen una garantía de disponibilidad de al menos el 99,99%.
- El nivel de hiperescala de Azure SQL Database con una réplica tiene una garantía de disponibilidad de al menos el 99,95% y el 99,9% para cero réplicas.
- El nivel crítico empresarial de Azure SQL Database configurado con replicación geográfica tiene una garantía de objetivo de punto de recuperación (RPO) de 5 segundos para el 100% de las horas implementadas.
- El nivel crítico empresarial de Azure SQL Database configurado con replicación geográfica tiene una garantía de objetivo de tiempo de recuperación (RTO) de 30 segundos para el 100% de las horas implementadas.

4. Amazon Aurora:

- Compromiso de servicio

AWS hará todos los esfuerzos comercialmente razonables para que los clústeres Multi-AZ estén disponibles con un porcentaje de tiempo de actividad mensual de al menos el 99,99% durante cualquier ciclo de facturación mensual (el "Compromiso de servicio"). En caso de que Amazon Aurora no cumpla con el Compromiso de servicio, será elegible para recibir un Crédito de servicio como se describe a continuación.

- Créditos de servicio

Los Créditos de servicio se calculan como un porcentaje de los cargos pagados por usted por los Clústeres Multi-AZ que no cumplieron con el Compromiso de servicio en un ciclo de facturación de acuerdo con el programa a continuación.

Monthly Uptime Percentage	Service Credit Percentage
Less than 99.99% but equal to or greater than 99.0%	10%
Less than 99.0% but equal to or greater than 95.0%	25%
Less than 95.0%	100%

Como garantizar un SLA similar a un servicio en la nube en infraestructura en premisas.

Es claro que antes de la existencia de los servicios en la nube eran necesarios los data centers en premisas por cuestiones que se creían de seguridad, fiabilidad, accesibilidad, tiempo de actividad, etc. Muchos de los beneficios declarados de los centros de datos locales tuvieron su legitimidad en algún tiempo. Pero desde el surgimiento de la “nube” estos supuestos beneficios de los centros de datos locales ya no son ciertos, ya que la nube es superior a las instalaciones para alojar la mayoría de las aplicaciones, los datos y servicios, es decir no se puede prestar un servicio similar porque simplemente ya no se comparan. Un SLA para un servicio en premisas podría garantizar rendimiento de primer nivel y el tiempo de actividad ofreciendo los siguientes servicios:

- Disponibilidad de Energía

- Temperatura
- Servicios de ancho de banda
- Permitir a los clientes solicitar un crédito en su fondo si una instalación no cumple con el compromiso de tiempo de actividad relevante.
- ITS proporcionará la infraestructura, tecnología, personas, procesos y herramientas de monitoreo necesarias para los servicios del centro de datos.

Bibliografía

<https://www.datacenterknowledge.com/machine-learning/nvidia-unveils-morpheus-ai-cybersecurity-framework>

<https://www.itprc.com/postgresql-monitoring/>

<https://aws.amazon.com/rds/aurora/sla/>

<https://www.logicata.com/blog/aws-service-level-agreement/>

<https://www.amazonaws.cn/en/rds/sla/ningxia/>

<https://www.oracle.com/cloud/iaas/sla.html>

<https://www.oracle.com/assets/paas-iaas-pub-cld-srvs-pillar-4021422.pdf>

<https://www.oracle.com/a/ocom/docs/oracle-cloud-infrastructure-sla-support.pdf>

<https://www.comparitech.com/net-admin/best-postgresql-monitoring-tools/>

<https://www.dnsstuff.com/mysql-tools>

https://www.manageengine.com/latam/applications_manager/mysql-management.html

<https://www.datadoghq.com/blog/postgresql-monitoring/#replication-and-reliability>

<https://severalnines.com/database-blog/monitoring-percona-distribution-postgresql-key-metrics>

<https://docs.oracle.com/es-ww/iaas/mysql-database/doc/mysql-database-metrics.html>

<https://docs.vmware.com/es/vRealize-Operations-Manager/8.3/com.vmware.vcom.core.doc/GUID-F43B8864-AE04-4B92-9BDE-BA03578EC7C1.html>