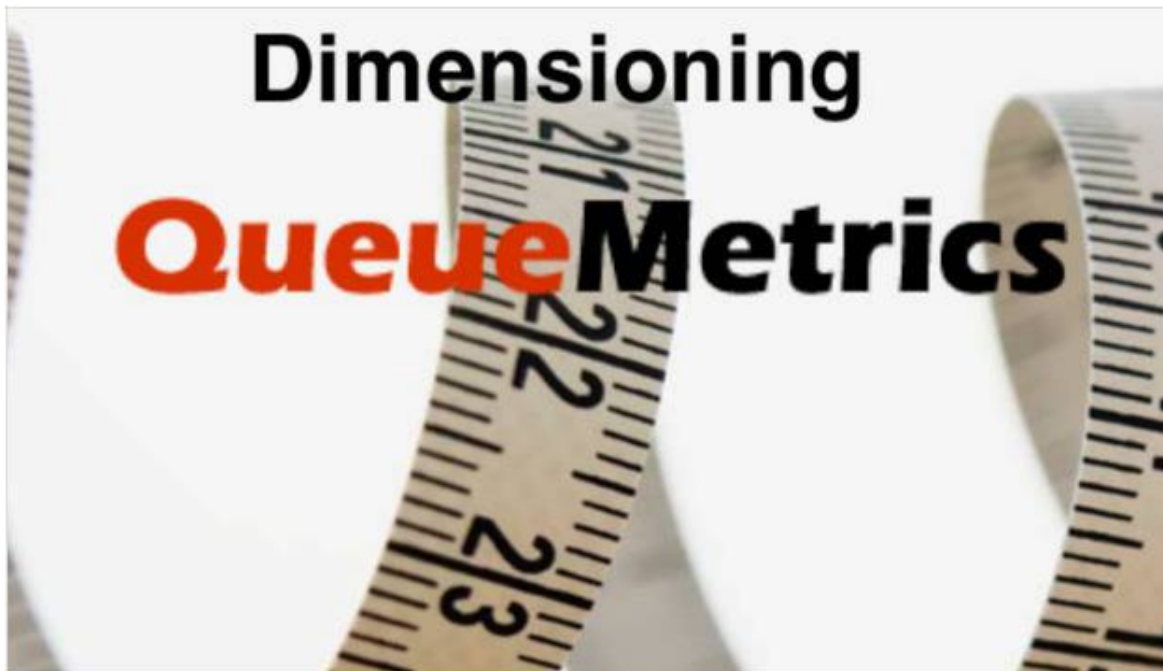


**Loway** presents

## Suite de dimensionamiento QueueMetrics Call Center



### Dimensionar un sistema QueueMetrics

QueueMetrics no es una aplicación especialmente “hambrienta” de recursos y está diseñada para ser bastante ahorrativa, pero al tratarse de una herramienta que se puede usar de varias maneras y para ejecutar informes posiblemente muy grandes, a veces es necesario hacer un pequeño análisis de antemano. Si bien los números reales dependen en gran medida de los patrones de uso específicos, existen algunas pautas generales que se pueden seguir para dimensionar.

Para dimensionar un sistema necesitarás:

- La cantidad de agentes que accederán al sistema simultáneamente.
- La cantidad de llamadas esperadas por día que planea analizar.
- Si va a utilizar llamadas al API para alimentar sistemas externos.

### Qué es QueueMetrics

QueueMetrics es un conjunto de informes y monitoreo altamente escalable que responde a las necesidades de miles de centros de contacto en todo el mundo. Ofrece una amplia gama de beneficios integrados como monitoreo de la productividad del agente, medición de objetivos, seguimiento de tasas de conversión, análisis de estadísticas de campañas en tiempo real y una interfaz fácil de usar. Está disponible On Premise o como un servicio de solución alojado en la Nube.

## **Uso del disco duro**

QueueMetrics realiza un seguimiento de los eventos de la cola, y puede esperar aproximadamente 5/10 eventos por llamada. Cada evento tendrá ~ 200 bytes en el disco, por lo que almacenar un millón de llamadas puede requerir ~ 2Gb en el disco en el peor de los casos. Este suele ser el elemento más grande en la base de datos QM.

Las grabaciones de audio y los documentos vinculados (por ejemplo, las transcripciones) se almacenan externamente y QueueMetrics los recuperará de un sistema de archivos conectado (mounted), por lo que no hay duplicación de datos.

El propio sistema QueueMetrics, incluida la máquina virtual Java y el motor de base de datos, requiere ~ 100 Mb en un sistema CentOS en blanco.

Como el rendimiento del disco es muy importante, cuanto más rápidos sean los discos (por ejemplo, SSD), mejor se ejecutará QM.

Como regla general, debe tener el doble del tamaño de la tabla de mayor dimensión como espacio libre en el disco, para permitir el mantenimiento (optimización de tablas) y el espacio temporal para copias de seguridad.

## **Uso de memoria**

QueueMetrics crea una imagen del conjunto de datos activos en la memoria, además de los datos que se requieren para los análisis que se solicitan. Para un sistema de 100 agentes con 4/5 supervisores, 1G a 2G son generalmente adecuados. Las llamadas a la API de informes cuentan como usuarios.

Los agentes, los tableros y las vistas en tiempo real mantienen en la memoria una imagen durante las últimas 24 horas, por lo general, de 10 a 20 Mb para cada usuario concurrente (el agente registrado, visualización de la Wallboard activa) es adecuado.

Por ejemplo, un sistema de 100 agentes con aproximadamente 50 usuarios registrados al mismo tiempo y 10 tableros abiertos puede requerir  $20M * 60 = 1.2G$  de memoria para actividades en tiempo real más 1G a 2G para informes = ~ 3.5 G disponibles.

Tener más memoria disponible ayudará al Garbage Collector a manejar los volúmenes máximos sin pausas, por lo que un heap de 4G Java podría ser adecuado. Si se dimensiona un servidor físico, Java usará ~ 200M para tareas generales, y el sistema operativo requerirá algo más de memoria como disco y caché de base de datos, por lo que 6G debería ser suficiente.

## **Uso de CPU**

En los sistemas modernos de múltiples núcleos, la CPU rara vez es un problema. Cada solicitud de QM se ejecuta en su propio subproceso, por lo que QM se amplía de forma natural en sistemas de

múltiples núcleos. Por lo general, no es un gran problema si una página o informe en tiempo real toma un segundo más o menos cuando se ejecuta, por lo que el rendimiento de la CPU no suele ser un cuello de botella.

Para dimensionar un sistema de 100 agentes con 50 agentes que inician sesión, se puede esperar que una página RT tome aproximadamente 50-100 ms de tiempo de CPU, por lo que serán adecuados de dos a cuatro núcleos.

## Consejos y trucos

- La virtualización suele ser una buena idea para los servidores QueueMetrics. Como QueueMetrics no es especialmente sensible a los tiempos precisos, ya que no maneja la voz en sí, funciona bien en la virtualización. Con la virtualización, puede tomar y restaurar fácilmente instantáneas y reasignar recursos (núcleos de CPU y memoria) según los patrones de uso reales.
- A menos que el sistema sea bastante pequeño (<50 agentes), sugerimos ejecutar QueueMetrics en un servidor virtual separado del PBX principal. Lo único que queda por instalar en el PBX es una pequeña herramienta de carga llamada uniloader que enviará datos de eventos de cola a QueueMetrics.
- El número de solicitudes simultáneas puede reducirse, si utiliza las API, reduciendo la tasa de pollig y/o ejecutando grandes informes analíticos cuando el sistema está inactivo, por ejemplo fuera de horas.
- La velocidad de actualización de las páginas en tiempo real se puede modificar. Convertir una actualización de 5 segundos en una actualización de 7 segundos reducirá la carga en un ~ 30% sin mucha diferencia para los usuarios.
- Ejecución de informes grandes cuando el sistema está inactivo: puede hacer que QueueMetrics genere informes automáticamente durante la noche y los envíe por correo electrónico.
- Medición: la JVM expone una gran cantidad de métricas a través de su interfaz JMX. Si ejecuta un servicio importante de QueueMetrics, le sugerimos configurar un motor de monitoreo para rastrear el número y la duración de los ciclos de Garbage Collection para que pueda ser rastreado en el tiempo. Como mínimo, en cualquier sistema QueueMetrics, debe buscar `java.lang.OutOfMemoryError` en los registros de QueueMetrics.
- Si tiene mucha memoria disponible, puede configurar la JVM para que use un recopilador de rendimiento, consulte, por ejemplo, <https://docs.oracle.com/javase/8/docs/technotes/guides/vm/gctuning/parallel.html>: para obtener el mejor rendimiento al precio de un mayor uso de memoria (~ 2x). Esto solo debe hacerse si está experimentando grandes pausas de recopilación, según lo medido por el

seguimiento de las estadísticas de JMX. De forma predeterminada, QueueMetrics utiliza el recopilador G1, que funciona bien en la mayoría de las cargas de trabajo sin ajuste explícito.

#### Referencias QueueMetrics:

Para obtener más información técnica sobre la solución de centro de llamadas de QueueMetrics, consulte el Manual del usuario.

Visite <https://www.queuemetrics.com/> para una prueba completa de 30 días.

Asista a nuestros seminarios web gratuitos para una demostración en vivo de QueueMetrics.

*Copyright ©Loway 2018 · all rights reserved*