

Manual de objetivos de nivel de servicio (SLO) para ejecutivos

Por qué los nuevos paradigmas de experiencia del cliente requieren que los ejecutivos se planteen objetivos de nivel de servicio



Resumen ejecutivo

Las aplicaciones modernas son más complejas y están más distribuidas que nunca, y su rendimiento también es más crítico para el negocio que nunca. Como resultado, los ejecutivos se sienten obligados a demostrar el impacto empresarial de su organización. Por suerte, las mismas tecnologías empleadas para modernizar y gestionar aplicaciones crean oportunidades para vincular directamente el impacto de la TI al negocio. Los equipos de TI y plataformas han encontrado nuevas maneras de medir la salud de sus entornos, fijar expectativas para sus aplicaciones y vincular sus esfuerzos al contexto empresarial, es decir, objetivos de nivel de servicio. Dada la importancia empresarial de las aplicaciones modernas, los grandes ejecutivos deben encauzar el planteamiento de la organización hacia la definición de los objetivos de nivel de servicio en términos de impacto empresarial y experiencia del cliente.

En este documento, repasaremos los enfoques y dificultades más habituales a la hora de definir e implementar objetivos de nivel de servicio (SLO), y ofreceremos asesoramiento en materia de buenas prácticas para obtener mejores resultados empresariales.

Contexto

A día de hoy, "la satisfacción del usuario final con el rendimiento y la fiabilidad de las aplicaciones es fundamental para el éxito de las operaciones empresariales digitales".¹ Ofrecer una experiencia de usuario final excelente, en cambio, resulta más complicado, ya que las aplicaciones modernas y las plataformas e infraestructuras en las que funcionan son más complejas y están más distribuidas que nunca. De hecho, como puede verse en nuestro informe sobre el estado del multicloud de 2021, se citó la complejidad como uno de los principales retos a la hora de alcanzar objetivos empresariales.²

Junto con el desarrollo de las aplicaciones modernas, "el paso a la infraestructura ágil (que incluye la TI híbrida, el multicloud y los contenedores)" ha hecho que los equipos de TI tengan que esforzarse sobremedida por encontrar sentido a cantidades enormes de datos y "ha puesto en peligro la viabilidad de las herramientas tradicionales de monitorización de infraestructuras".³

En respuesta a estos retos, los equipos de TI y plataformas han encontrado nuevas maneras de medir la salud de sus entornos, fijar expectativas para sus aplicaciones y vincular sus esfuerzos al impacto empresarial. Las expectativas de los usuarios finales son cada vez más altas y, por tanto, es fundamental que los ejecutivos se aseguren de que sus organizaciones emplean métricas significativas a la hora de evaluar el rendimiento de sus aplicaciones y su impacto en el negocio.

1 Fuente: Cuotas de mercado mundial del software de gestión del rendimiento de las aplicaciones, 2020 | Junio de 2021, IDC #US47989021

2 Fuente: Informe sobre el estado del multicloud de Turbonomic de 2021, Encuesta CNCF de 2020

3 Fuente: Monitoring and Observability for Modern Services and Infrastructure | Junio de 2020, Gartner G00720854

El enfoque actual: laborioso y reactivo

Un enfoque habitual que los equipos de TI y plataformas suelen adoptar para medir la salud de sus entornos es el de identificar y configurar indicadores de nivel de servicio, acuerdos de nivel de servicio y objetivos de nivel de servicio.

- **Un indicador de nivel de servicio (SLI) es la métrica empleada por los equipos para medir el rendimiento de un servicio en relación con un objetivo.**
- **Un acuerdo de nivel de servicio (SLA) es una promesa de disponibilidad que los equipos de TI y plataformas hacen a los clientes o usuarios finales.**
- **Un objetivo de nivel de servicio (SLO) es la meta u objetivo de nivel de servicio, medido con un SLI, que los equipos se comprometen a alcanzar.**

Los SLO son especialmente importantes porque definen los criterios y fijan unas expectativas claras en cuanto a rendimiento de las aplicaciones de una organización.

Actualmente, la mayoría de organizaciones pasan por un proceso largo y laborioso de configuración manual de los SLO. Los equipos de TI deben empezar por determinar qué servicios de una aplicación afectan directamente al negocio y la experiencia del usuario final. A continuación, deben identificar qué métricas usar como indicadores de nivel de servicio. Las métricas SLI más habituales son la disponibilidad, la latencia y el rendimiento de transacciones. Una vez elegidos los SLI más adecuados, los equipos de TI deben determinar sus objetivos de SLO para esa métrica y un periodo de medición concreto. Tras completar este proceso, deben crear estimaciones de errores para cada SLO y vincularlas a un sistema de alertas basado en umbrales. Muchas organizaciones adoptan este enfoque basado en umbrales porque al ser humano le resulta imposible monitorizar el rendimiento de las aplicaciones las 24 horas del día todos los días.

Por desgracia, fijar umbrales de SLO no soluciona las cuestiones de rendimiento. Esta estrategia de TI no es eficaz porque resulta demasiado reactiva para las infraestructuras ágiles de hoy, con aplicaciones que se ejecutan en entornos multicloud o contenerizados. Si algo sale mal con un servicio y se dispara un umbral o una alerta, ya se habrá producido una degradación del rendimiento de ese servicio y habrá dado como resultado una experiencia del usuario final insatisfactoria.

Los equipos de TI han tratado de mejorar ese sistema de umbrales/alertas implementando el escalado automático horizontal de pods (HPA) para las aplicaciones que se ejecutan en entornos contenerizados. Sin embargo, este sistema tampoco garantiza una buena experiencia del usuario final ni impide la degradación del rendimiento. A semejanza del proceso de configuración de SLO y con el fin de establecer el escalado automático horizontal para satisfacer las demandas de recursos, los equipos de TI deben identificar métricas que representen mejor las necesidades de recursos, configurar objetivos, fijar umbrales y hacer pruebas. Este proceso debe repetirse para cada servicio de una aplicación. Algunas aplicaciones ofrecen cientos de servicios diferentes, por lo que implementar HPA a escala resulta muy complicado. Además, las distintas políticas de HPA siguen dependiendo de umbrales y no se correlacionan ni cumplen con las demás, con lo que escalar un servicio podría afectar negativamente a otro. Por último, no se trata de un ejercicio aislado, y las políticas de escalado de HPA requieren de reconfiguración y monitorización continuas para ser eficaces.



Cómo pensar en los SLO...

Defina los SLI y los SLO en términos de experiencia del cliente.

Según expertos del sector, "centrarse en los sistemas de gestión y las aplicaciones para optimizar la experiencia del usuario final es una prioridad importante, **ya que un rendimiento rápido y un tiempo de actividad del 100 % son elementos esenciales para el éxito del negocio digital**".⁴ Puesto que un rendimiento rápido y un tiempo de actividad permanente se han convertido en elementos esenciales, las organizaciones no pueden perder el tiempo recopilando datos que no sean indicativos de la experiencia del usuario final a la hora de evaluar el rendimiento de las aplicaciones.

Recopilar los datos adecuados no es tarea fácil. Por ejemplo, aunque la disponibilidad es una métrica de uso habitual, no es una extensión directa del rendimiento, porque una aplicación puede estar disponible y, aun así, experimentar cuellos de botella en los recursos y una degradación del rendimiento. Hay muchas métricas que las organizaciones pueden emplear para medir el rendimiento, por lo que es importante que la TI y la línea de negocio acuerden qué datos recopilar y reflejar en los informes. Tal y como afirma Gartner, "es fundamental elegir SLI representativos y valiosos. En la mayoría de los casos, es posible que una métrica basada en infraestructuras ("memoria disponible" o "porcentaje de nodos de trabajador libres") no sea significativa porque no suele ser algo que les importe a los usuarios de un servicio. Elija SLI que sean mediciones directas de la experiencia que los usuarios tienen con su servicio".⁵

En pocas palabras, los SLO deben usarse como una herramienta que indica a las organizaciones si sus aplicaciones están haciendo lo que tienen que hacer por su negocio. Si las aplicaciones no funcionan como deberían, las organizaciones que cuentan con un buen SLO definido en el nivel adecuado de la aplicación deberían saber exactamente qué medidas tendrán más impacto para recuperar el rendimiento. Las distintas organizaciones necesitan medir distintas métricas específicas de sus líneas de negocio. Sin embargo, a medida que las organizaciones se van alejando de las arquitecturas de aplicaciones monolíticas para ser más modernas y estar más distribuidas, los indicadores de rendimiento tradicionales, como la alta memoria o el uso de la CPU, se vuelven menos útiles. En lugar de eso, las organizaciones deben tratar de definir SLI/SLO en términos de métricas que suelen ser relevantes para el negocio, como el tiempo de respuesta o el rendimiento de transacciones. Estas métricas son una medida más directa del rendimiento porque, por ejemplo, al definir SLO para el rendimiento de transacciones, los equipos de TI pueden saber exactamente cómo se está atendiendo a las solicitudes de cada pod o máquina virtual. Esta métrica supone una evaluación más directa del rendimiento que definir SLO basados en métricas de utilización y asociarlos a cosas como el tiempo de respuesta o el rendimiento de transacciones.

⁴ Fuente: Cuotas de mercado mundial del software de gestión del rendimiento de las aplicaciones, 2020 | Junio de 2021, IDC #US47989021

⁵ Solution Path for Modern Infrastructure and Application Monitoring | Junio de 2019, Gartner

Pese a que el tiempo de respuesta o el rendimiento de transacciones son formas muy habituales de evaluar la experiencia del cliente y el rendimiento de las aplicaciones, estas métricas no funcionan para todas las organizaciones. Por ejemplo, un proveedor de servicios de infraestructura de escritorio virtual no definiría SLO para el rendimiento de transacciones, sino para incidencias de soporte. Las incidencias de soporte serían una métrica de mayor impacto para un SLO porque se trata de una medida directa del rendimiento de los escritorios virtuales. En última instancia, los SLI/SLO deben definirse específicamente para cada organización y su línea de negocio.

Se requiere un análisis continuo de la demanda cambiante de aplicaciones.

Las aplicaciones e infraestructuras modernas son elásticas, y la demanda de recursos es dinámica. Debido a este cambio constante, entender las relaciones entre las distintas fuentes y tipos de datos resulta muy complicado. Las organizaciones suelen usar diversas herramientas para monitorizar las distintas capas de la pila en los distintos equipos, y siempre tratando de resolver distintos aspectos del mismo problema. Por consiguiente, esta falta de coordinación desemboca en muchas pistas falsas sobre la causa principal de un problema. Asimismo, los datos deben recopilarse de forma continua porque, de no ser así, los equipos de TI y plataformas tendrán dudas de qué datos necesitan para solucionar un problema determinado. Tal y como indica Gartner en un informe de monitorización y observabilidad de 2020, "la recopilación de los datos sobre la posible causa debe realizarse de forma continua, ya que podría pasarse por alto la causa real si se hace en respuesta a un síntoma".⁶

Con el fin de evitar este dilema, las organizaciones necesitan un repositorio central capaz de agregar y correlacionar datos de cada capa de una pila de aplicaciones. Con este sistema, pueden reunir todos los datos que recogen de forma continua y depurarlos hasta que se ajusten a sus SLO. Al hacer este análisis continuo, las organizaciones pueden contextualizarlo todo y saber qué objetivo deben alcanzar para un rendimiento satisfactorio de las aplicaciones. Y si este objetivo no se cumple, pueden saber cuáles son los problemas en las capas subyacentes de la pila de aplicaciones que hay que resolver. En última instancia, implementar sistemas que puedan hacer este análisis continuo es la única manera de estar a la altura de las necesidades de recursos dinámicos de las aplicaciones e infraestructuras modernas.



Automatice los recursos dinámicos.

Al igual que la naturaleza dinámica de las aplicaciones modernas y de la infraestructura en la que se ejecutan requiere de un análisis continuo, los ejecutivos y sus organizaciones también deben tratar de implementar una automatización continua si pretenden aprovechar al máximo los SLO y crear un sistema preventivo para gestionar sus aplicaciones. En un informe de análisis reciente de IDC, estos afirmaron que para seguir siendo competitivos en el futuro deben "tener en cuenta el papel de la automatización en las prestaciones de un producto".⁷

⁶ Fuente: Monitoring and Observability for Modern Services and Infrastructure | Junio de 2020, Gartner G00720854

⁷ Fuente: Cuotas de mercado mundial de software de gestión del rendimiento de las aplicaciones, 2020 | Junio de 2021, IDC #US47989021

El desarrollo de las infraestructuras modernas prometía la resiliencia y elasticidad de las aplicaciones en el futuro, pero muchas organizaciones luchan por mantener el rendimiento de las suyas hoy en día. Implementar la automatización en la gestión de las aplicaciones modernas es fundamental porque es necesario para cosechar los beneficios de la resiliencia y la elasticidad. Sin la automatización, las resoluciones que son posibles mediante la configuración de los SLO adecuados y la recopilación continua de datos no son eficaces. Los equipos de TI y plataformas no pueden garantizar el rendimiento de las aplicaciones sin automatizar los recursos dinámicos: para cuando se active una alerta y se ejecute manualmente una decisión relativa a los recursos, ya se habrá producido una degradación del rendimiento.

En última instancia, el problema con este enfoque es que se basa en el hecho de que ya se ha producido un problema.

La automatización puede ir más allá de una simple reacción a un evento activado por un umbral. Al aprovechar la recopilación continua de datos e identificar SLI y SLO claves para su línea de negocio, las organizaciones pueden hacer que el software genere decisiones factibles que deberían automatizarse. Si, como organización, toma las medidas necesarias y se compromete con la automatización, puede crear un entorno realmente elástico que gestione de forma proactiva los recursos de sus aplicaciones, garantizando de este modo el rendimiento continuo. Este tipo de automatización requiere de un sistema inteligente capaz de analizar un entorno que cambia de manera dinámica y de automatizar las decisiones en cadena necesarias, de forma que los problemas se solucionen *antes* de que se produzca una degradación del rendimiento. Resulta imposible lograr este tipo de rendimiento y elasticidad mediante un sistema de hojas de cálculo y alertas.

Para poder dejar atrás el enfoque reactivo, debe involucrarse a todas las partes interesadas, tanto de los equipos que gestionan las aplicaciones como de los que gestionan la infraestructura. A los propietarios de aplicaciones y productos les suele costar ceder el control de sus aplicaciones, y esta reticencia se debe a una falta de confianza en la automatización. Pero puede superarse. La automatización requiere una transformación cultural de su organización de TI. Para lograr verdadera elasticidad y resiliencia, las organizaciones deben ser capaces de confiar en las acciones que van a automatizarse. Con SLO relevantes basados en el negocio y las aplicaciones y directamente vinculados a los recursos dinámicos de la plataforma y la infraestructura resulta más fácil para los propietarios de aplicaciones y productos confiar en la automatización y sentirse cómodos con la posibilidad de que esta se implemente totalmente en los flujos de trabajo empresariales.

Cosas que tener en cuenta

Las aplicaciones se volverán cada vez más complejas a medida que las organizaciones vayan adoptando infraestructuras ágiles, como la TI híbrida, las plataformas de contenedores y el multicloud. Con estos avances futuros, las organizaciones que más prosperarán serán aquellas que satisfagan grandes expectativas en cuanto a experiencia del cliente y rendimiento de las aplicaciones. Las mejores prácticas consisten en definir los SLO en términos de experiencia del cliente, analizar continuamente la demanda cambiante de aplicaciones y automatizar los recursos dinámicos. Si desean alcanzar el éxito en los negocios digitales, las organizaciones deben dar los pasos necesarios para implementar una automatización que asigne recursos dinámicos a las aplicaciones con el fin de satisfacer la demanda cambiante y los SLO de la empresa.

Garantice automáticamente los SLO de sus aplicaciones con Turbonomic

Turbonomic convierte los datos en acción y ofrece una automatización que evita los riesgos para el rendimiento de las aplicaciones al tiempo que maximiza la elasticidad.

Modernizar aplicaciones e infraestructuras fundamentales para el negocio es una inversión con muchos beneficios, pero para lograr elasticidad, resiliencia y rapidez de comercialización, se necesita un software que analice continuamente su entorno y ejecute las decisiones correctas de asignación de recursos en el momento adecuado para garantizar el rendimiento de las aplicaciones. Con Turbonomic, puede correlacionar el tiempo de respuesta de las aplicaciones, el rendimiento de transacciones u otros SLI/SLO para asignar recursos de forma dinámica. Cuando la demanda fluctúa, la asignación de recursos dinámicos de Turbonomic garantiza el rendimiento continuo de las aplicaciones.


Un HPA no da la talla. Turbonomic usa un análisis descendente de pila completa para garantizar dinámicamente sus SLO. Usted establece sus SLO y nuestro software con IA garantiza que la plataforma y la infraestructura subyacente ofrezcan los recursos que necesitan para alcanzar esos SLO, independientemente de dónde se ejecuten sus aplicaciones.

Integración sin problemas en los flujos de trabajo empresariales. Gracias a la integración de Turbonomic con webhooks, podrá inyectar fácilmente las acciones de Turbonomic en los ciclos de vida de las aplicaciones, las canalizaciones de DevOps e infraestructuras, los flujos de trabajo de aprobaciones y auditoría y los procesos de comunicación.

Reduzca al mínimo el trabajo manual: los desarrolladores, DevOps y SRE no necesitarán establecer umbrales, restricciones o políticas de escalado automático. El software toma por ellos las decisiones de recursos adecuadas y ofrece acciones que pueden automatizarse.

No gaste de más en capacidad: ya no tendrá que depender de los desarrolladores para tomar las decisiones adecuadas en materia de recursos (¡suelen suministrar en exceso para curarse en salud!) Nuestro software determina exactamente qué recursos necesitan los servicios, todos ellos basados en la demanda de aplicaciones.

Planifique el crecimiento de forma fácil y rápida: simule la incorporación de nuevos servicios con nuestro software. Determine exactamente lo que necesita para apuntalar el crecimiento.



Pruebe
Turbonomic
hoy mismo

turbonomic.com/try-SLO

© Copyright IBM Corporation 2022

IBM España, S.A.

Santa Hortensia, 26-28

28002 Madrid

Producido en los Estados Unidos de América

Enero de 2022

IBM y el logotipo de IBM son marcas comerciales o marcas registradas de International Business Machines Corporation, en Estados Unidos o en otros países. Los demás nombres de productos y servicios pueden ser marcas comerciales de IBM u otras empresas. Una lista actual de marcas comerciales de IBM está disponible en ibm.com/trademark.

Este documento está actualizado en la fecha inicial de publicación e IBM puede modificarlo en cualquier momento. No todas las ofertas están disponibles en todos los países en los que opera IBM.

Los datos de rendimiento y ejemplos de clientes mencionados se presentan únicamente con fines ilustrativos. Los datos reales de rendimiento pueden variar en función de las configuraciones y condiciones de funcionamiento específicas.

LA INFORMACIÓN DE ESTE DOCUMENTO SE OFRECE "TAL CUAL ESTÁ" SIN NINGUNA GARANTÍA, NI EXPLÍCITA NI IMPLÍCITA, INCLUIDAS, ENTRE OTRAS, LAS GARANTÍAS DE COMERCIALIZACIÓN, ADECUACIÓN A UN FIN CONCRETO Y CUALQUIER GARANTÍA O CONDICIÓN DE INEXISTENCIA DE INFRACCIÓN.

Los productos de IBM están garantizados según los términos y condiciones de los acuerdos bajo los que se proporcionan.

