

Distributed Message Service for Kafka

Descripción general del servicio

Edición 01
Fecha 2022-11-08




Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2022. Todos los derechos reservados.

Quedan terminantemente prohibidas la reproducción y/o la divulgación totales y/o parciales del presente documento de cualquier forma y/o por cualquier medio sin la previa autorización por escrito de Huawei Cloud Computing Technologies Co., Ltd.

Marcas registradas y permisos



El logotipo  y otras marcas registradas de Huawei pertenecen a Huawei Technologies Co., Ltd. Todas las demás marcas registradas y los otros nombres comerciales mencionados en este documento son propiedad de sus respectivos titulares.

Aviso

Es posible que la totalidad o parte de los productos, las funcionalidades y/o los servicios que figuran en el presente documento no se encuentren dentro del alcance de un contrato vigente entre Huawei Cloud y el cliente. Las funcionalidades, los productos y los servicios adquiridos se limitan a los estipulados en el respectivo contrato. A menos que un contrato especifique lo contrario, ninguna de las afirmaciones, informaciones ni recomendaciones contenidas en el presente documento constituye garantía alguna, ni expresa ni implícita.

Huawei está permanentemente preocupada por la calidad de los contenidos de este documento; sin embargo, ninguna declaración, información ni recomendación aquí contenida constituye garantía alguna, ni expresa ni implícita. La información contenida en este documento se encuentra sujeta a cambios sin previo aviso.

Índice

1 ¿Qué es DMS for Kafka?.....	1
2 Ventajas del producto.....	2
3 Escenarios de aplicación.....	4
4 Especificaciones.....	7
5 Comparación de Kafka y RabbitMQ.....	12
6 Comparación de DMS for Kafka y Kafka de código abierto.....	15
7 Notas y restricciones.....	18
8 Servicios relacionados.....	20
9 Conceptos básicos.....	21
10 Gestión de permisos.....	23
11 Facturación.....	26

1 ¿Qué es DMS for Kafka?

Apache Kafka es un middleware de mensajes distribuidos que cuenta con alto rendimiento, persistencia de datos, escalabilidad horizontal y procesamiento de datos de transmisión. Adopta el patrón de publicación y suscripción y se utiliza ampliamente para la recopilación de registros, transmisión de datos, análisis de sistemas en línea/fuera de línea y monitoreo en tiempo real.

El servicio de mensajes distribuidos (DMS) for Kafka es un servicio de espera de mensajes basado en Apache Kafka. Este servicio proporciona instancias premium de Kafka. Los recursos de cómputo, almacenamiento y ancho de banda utilizados por una instancia son ocupados exclusivamente por el usuario. Puede solicitar instancias según sea necesario y personalizar particiones y réplicas para los temas de las instancias. Las instancias se pueden usar de inmediato, lo que elimina la presión de implementación y operación y mantenimiento para que pueda concentrarse en desarrollar sus servicios.

Guía del lector

Esta documentación presenta DMS for Kafka y sus diferencias con Apache Kafka. Obtendrá información detallada sobre las especificaciones, las operaciones de la consola, las llamadas a la API y el acceso del cliente a las instancias de Huawei Cloud DMS for Kafka.

Para obtener más información sobre los conocimientos básicos de Kafka o detalles técnicos sobre la creación y recuperación de mensajes, visite el [sitio web oficial de Apache Kafka](#).

2 Ventajas del producto

Huawei Cloud DMS for Kafka proporciona colas de mensajes fáciles de usar basadas en Apache Kafka. Los servicios se pueden migrar rápido a la nube sin ningún cambio, lo que reduce los costos de mantenimiento y uso.

- **Despliegue rápido**

Simplemente configure la información de la instancia en el DMS para la consola de Kafka, envíe su pedido, se creará e implementará automáticamente una instancia completa de Kafka.

- **Migración de servicios sin modificaciones**

DMS for Kafka es compatible con las API de Kafka de código abierto y soporta todas las funciones de procesamiento de mensajes de Kafka de código abierto.

Si sus servicios de aplicación se desarrollan basados en Kafka de código abierto, puede migrarlos fácilmente a Huawei Cloud DMS for Kafka después de especificar algunas configuraciones de autenticación.

 **NOTA**

Las instancias de Kafka son compatibles con Apache Kafka v1.1.0, v2.3.0 y v2.7. Mantenga las versiones del cliente y las del servidor iguales.

- **Seguridad**

Las operaciones en las instancias de Kafka se registran y se pueden auditar. Los mensajes se pueden cifrar antes del almacenamiento.

Además de la convalidación de Autenticación Simple y Capa de Seguridad (SASL), Virtual Private Clouds (VPC) y los grupos de seguridad también proporcionan controles de seguridad en el acceso a la red.

- **Confiabilidad de los datos**

Las instancias de Kafka admiten la persistencia y replicación de datos. Los mensajes se pueden replicar de forma síncrona o asíncrona entre réplicas.

- **Alta disponibilidad**

Kafka se ejecuta en clústeres, lo que permite la conmutación por error y la tolerancia a fallos para que los servicios puedan ejecutarse sin problemas.

Los corredores de una instancia de Kafka se pueden implementar en todas las AZ para mejorar aún más la disponibilidad del servicio.

- **O&M sencillos**

Huawei Cloud proporciona un conjunto completo de servicios de monitoreo y alarma, eliminando la necesidad de asistencia 24/7. Las métricas de instancia de Kafka se supervisan e informan, incluido el número de particiones, temas y mensajes acumulados. Puede configurar reglas de alarma y recibir notificaciones por SMS o correo electrónico sobre cómo se ejecutan sus servicios en tiempo real.

- Acumulación masiva y escalado automático

Kafka presenta una gran escalabilidad porque se ejecuta en un sistema distribuido o clúster. Puede configurar hasta 100 particiones para un tema. El espacio de almacenamiento también se puede ampliar. Esto significa que se pueden acumular miles de millones de mensajes, adecuados para escenarios que requieren alta simultaneidad, alto rendimiento y acceso a gran escala.

- Especificaciones flexibles

Puede personalizar el ancho de banda y el espacio de almacenamiento de la instancia y el número de particiones y réplicas de los temas de la instancia.

3 Escenarios de aplicación

Kafka es un popular middleware orientado a mensajes que ofrece una entrega de mensajes asíncronos y altamente confiable. Es ampliamente utilizado para la transmisión de datos entre diferentes sistemas en la aplicación empresarial, el pago, las telecomunicaciones, el comercio electrónico, las redes sociales, la mensajería instantánea, el vídeo, el Internet de las cosas y el Internet de las industrias de vehículos.

Comunicación asíncrona

Los mensajes no básicos o menos importantes se envían de forma asíncrona a los sistemas receptores, de modo que el proceso de servicio principal no se mantiene esperando los resultados de otros sistemas, lo que permite respuestas más rápidas.

Por ejemplo, Kafka se puede utilizar para enviar un correo electrónico de notificación y un mensaje SMS después de que un usuario se haya registrado en un sitio web, proporcionando respuestas rápidas durante todo el proceso de registro.

Figura 3-1 Registro de serie y notificación



Figura 3-2 Registro asíncrono y notificación por colas de mensajes



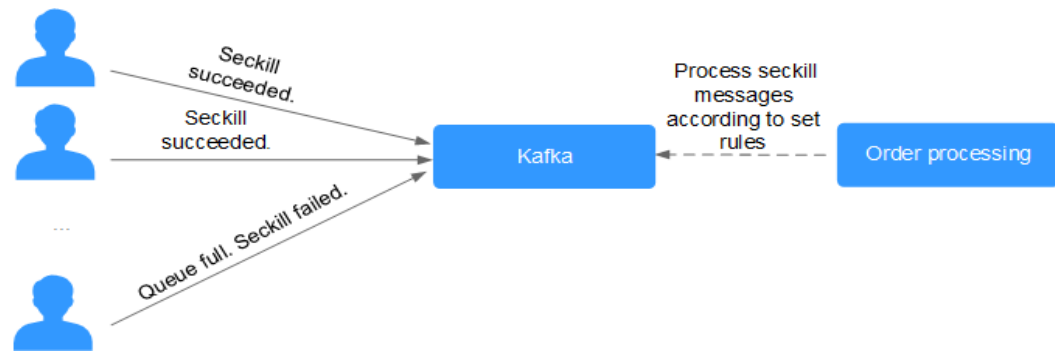
Control de tráfico

En los sistemas de comercio electrónico o en los sitios web a gran escala, existe una brecha de capacidad de procesamiento entre los sistemas anteriores y posteriores. Las ráfagas de tráfico de sistemas ascendentes con altas capacidades de procesamiento pueden tener un gran impacto en sistemas descendentes con menores capacidades de procesamiento. Por ejemplo, las promociones de ventas online implican una gran cantidad de tráfico que inunda los

sistemas de comercio electrónico. Kafka proporciona un búfer de tres días de forma predeterminada para cientos de millones de mensajes, como pedidos y otra información. De esta manera, los sistemas de consumo de mensajes pueden procesar los mensajes durante periodos fuera de pico.

Además, las ráfagas de tráfico de ventas flash que se originan en los sistemas frontend se pueden manejar con Kafka, lo que evita que los sistemas backend se bloqueen.

Figura 3-3 Manejo de ráfagas de tráfico con Kafka



Sincronización de registros

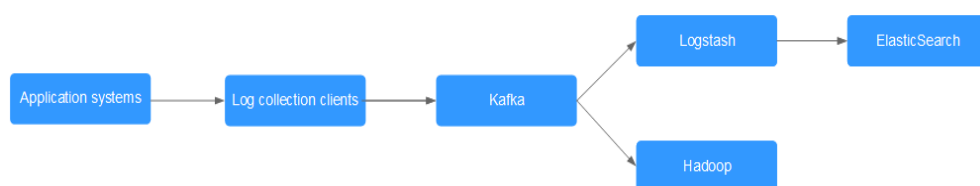
En los sistemas de servicio a gran escala, se recopilan registros de diferentes aplicaciones para la resolución rápida de problemas, el seguimiento de enlaces completos y la supervisión en tiempo real.

Kafka está diseñado originalmente para este escenario. Las aplicaciones asincrónicamente envían los mensajes de registro a las colas de mensajes a través de canales de transmisión fiables. Otros componentes pueden leer los mensajes de registro de las colas de mensajes para su análisis posterior, ya sea en tiempo real o fuera de línea. Además, Kafka puede recopilar información clave de registro para monitorear aplicaciones.

La sincronización de registros implica tres componentes principales: clientes de recopilación de registros, Kafka y aplicaciones de procesamiento de registros de back-end.

1. Los clientes de recopilación de registros recogen los datos de registro de un servicio de aplicación de usuario y los envían de forma asíncrona por lotes a los clientes Kafka. Los clientes Kafka reciben y comprimen mensajes por lotes. Esto solo tiene un impacto menor en el rendimiento del servicio.
2. Kafka persiste los registros.
3. Las aplicaciones de procesamiento de registros, como Logstash, se suscriben a mensajes en Kafka y recuperan mensajes de registro de Kafka. A continuación, los mensajes son buscados por los servicios de búsqueda de archivos o entregados a aplicaciones de big data como Hadoop para el almacenamiento y el análisis.

Figura 3-4 Proceso de sincronización de registros



 **NOTA**

Logstash es para el análisis de registros, el ElasticSearch es para la búsqueda de registros y el Hadoop es para el análisis de big data. Todas son herramientas de código abierto.

4 Especificaciones

Especificaciones de instancia de Kafka

Las instancias de Kafka son compatibles con Kafka v1.1.0, v2.3.0 y v2.7 de código abierto. Hay dos tipos de especificaciones de instancia. En las regiones CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg, las especificaciones de la instancia están representadas por la variante de ECS y el número de corredores. Las opciones disponibles son c6.2u4g.cluster, c6.4u8g.cluster, c6.8u16g.cluster, c6.12u24g.cluster y c6.16u32g.cluster. En otras regiones, las especificaciones de instancia están representadas por el ancho de banda de la instancia. Las opciones disponibles son 100 MB/s, 300 MB/s, 600 MB/s y 1200 MB/s.

NOTA

- El número de corredores varía según los recursos subyacentes, y los recursos subyacentes varían de una región a otra. En la siguiente tabla se enumeran las especificaciones.
- En la siguiente tabla, las transacciones por segundo (TPS) se calculan suponiendo que el tamaño de un mensaje es de 1 KB.
- v2.7 está disponible solo en las regiones CN East-Shanghai1, CN South-Guangzhou, y CN-Hong Kong.

Tabla 4-1 Especificaciones de instancia de Kafka

Ancho de banda	Corredores	Tipo de recursos subyacentes	Tipo de E/S	TPS (alto rendimiento)	TPS (Replicación Sincrónica)	Cantidad máxima de particiones
100 MB/s	3	c6_2 vCPUs 4 GB	Capacidad alta de E/S	100,000	60,000	300
		c6_2 vCPUs 4 GB	Capacidad ultraalta de E/S	100,000	80,000	300

Ancho de banda	Corredores	Tipo de recursos subyacentes	Tipo de E/S	TPS (alto rendimiento)	TPS (Replicación Sincrónica)	Cantidad máxima de particiones
300 MB/s	3	c6_4 vCPUs 8 GB	Capacidad alta de E/S	300,000	150,000	900
		c6_4 vCPUs 8 GB	Capacidad ultra alta de E/S	300,000	200,000	900
600 MB/s	4	c6_8 vCPUs 16 GB	Capacidad ultra alta de E/S	600,000	300,000	1800
1200 MB/s	8	c6_8 vCPUs 16 GB	Capacidad ultra alta de E/S	1.2 millones	400,000	1800

Las especificaciones enumeradas en [Tabla 4-1](#) están disponibles en regiones distintas de CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg.

Tabla 4-2 Especificaciones de instancia de Kafka

Variante	Corredores	TPS máximo por corredor	Particiones máximas por corredor	Máximo de grupos de consumidores por corredor	Máximo de conexiones de cliente por corredor	Espacio de almacenamiento
c6.2u4g.cluster	3 - 30	30,000	250	20	1000	600 GB–300,000 GB
c6.4u8g.cluster	3 - 30	100,000	500	100	2000	1200 GB–600,000 GB
c6.8u16g.cluster	3 - 30	150,000	1000	150	3000	2400 GB–900,000 GB
c6.12u24g.cluster	3 - 30	200,000	1500	200	4000	3600 GB - 900,000 GB

Variante	Corredores	TPS máximo por corredor	Particiones máximas por corredor	Máximo de grupos de consumidores por corredor	Máximo de conexiones de cliente por corredor	Espacio de almacenamiento
c6.16u32g.cluster	3 - 30	250,000	2000	200	5000	4800 GB–900,000 GB

Las especificaciones enumeradas en [Tabla 4-2](#) están disponibles en regiones CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg.

Selección de ancho de banda

El ancho de banda de una instancia de Kafka se refiere al ancho de banda máximo de lectura o escritura. Se recomienda seleccionar un ancho de banda 30% más alto que el requerido.

- 100 MB/s
Recomendada para hasta 3000 conexiones de clientes, 60 grupos de consumidores y 70 MB/s de tráfico de servicio.
- 300 MB/s
Recomendada para hasta 10,000 conexiones de clientes, 300 grupos de consumidores y 210 MB/s de tráfico de servicio.
- 600 MB/s
Recomendada para hasta 20,000 conexiones de clientes, 600 grupos de consumidores y 420 MB/s de tráfico de servicio.
- 1200 MB/s
Recomendada para hasta 20,000 conexiones de clientes, 600 grupos de consumidores y 840 MB/s de tráfico de servicio.

Selección de variante

- c6.2u4g.cluster con 3 corredores
Recomendado para hasta 3000 conexiones de clientes, 60 grupos de consumidores y TPS 100,000
- c6.4u8g.cluster con 3 corredores
Recomendado para hasta 10,000 conexiones de clientes, 300 grupos de consumidores y TPS 300,000
- c6.8u16g.cluster con 3 corredores
Recomendado para hasta 20,000 conexiones de clientes, 600 grupos de consumidores y TPS 600,000
- c6.12u24g.cluster con 3 corredores
Recomendado para hasta 20,000 conexiones de clientes, 600 grupos de consumidores y TPS 900,000

- c6.16u32g.cluster con 3 corredores
Recomendado para hasta 20,000 conexiones de clientes, 600 grupos de consumidores y TPS 1,200,000

Selección de espacio de almacenamiento

Las instancias de Kafka admiten almacenamiento con 1 hasta 3 réplicas. El espacio de almacenamiento es el espacio que consumen todas las réplicas. Al crear una instancia, especifique su espacio de almacenamiento en base al tamaño esperado del mensaje de servicio y el número de réplicas.

Por ejemplo, si el tamaño estimado del mensaje es de 100 GB, la capacidad del disco debe ser como mínimo: 100 GB x cantidad de réplicas + 100 GB (reservado).

El espacio de almacenamiento se puede ampliar a medida que su servicio crece.

Cantidad del tema

No hay límite en la cantidad de temas, pero hay un límite superior en el número agregado de particiones en los temas. Cuando se alcanza el límite de cantidad de partición, ya no se pueden crear temas.

El número de temas está relacionado con el número máximo de particiones permitidas (ver [Figura 4-1](#)) y el número especificado de particiones en cada tema (ver [Tabla 4-1](#) y [Tabla 4-2](#)).

Figura 4-1 Establecimiento del número de particiones

Create Topic

* Topic Name

Partitions Value range: 1 to 100

Replicas Value range: 1 to 3

Aging Time (h) Value range: 1 to 168

Synchronous Replication

Synchronous Flushing

OK

El número máximo de particiones para una instancia de 100 MB/s es de 300.

- Si el número de particiones de cada tema en la instancia es 3, el número de temas es $300/3 = 100$.
- Si el número de particiones de cada tema en la instancia es 1, el número de temas es $300/1 = 300$.

El número máximo de particiones permitidas para una instancia con `c6.2u4g.cluster` y 3 corredores es 750.

- Si el número de particiones de cada tema en la instancia es 3, el número de temas es $750/3 = 250$.
- Si el número de particiones de cada tema en la instancia es 1, el número de temas es $750/1 = 750$.

5 Comparación de Kafka y RabbitMQ

Kafka se basa en la pull y proporciona un mayor rendimiento. Es adecuado para recopilar y entregar grandes volúmenes de datos, como la recopilación y el análisis de registros. RabbitMQ no proporciona un rendimiento tan alto como Kafka, pero ofrece más funciones de cola de mensajes.

El siguiente es un análisis comparativo sobre el rendimiento, la fiabilidad de los datos, la disponibilidad del servicio y las funciones de Kafka y RabbitMQ.

Rendimiento

El rendimiento del middleware orientado a mensajes se mide por el rendimiento. Mientras que RabbitMQ proporciona decenas de miles de QPS, Kafka proporciona millones.

Sin embargo, si la idempotencia y las transacciones están habilitadas para Kafka, su rendimiento se verá comprometido.

Confiabilidad de los datos

Tanto Kafka como RabbitMQ proporcionan el mecanismo de replicación para garantizar una alta confiabilidad de los datos.

Disponibilidad del servicio

Kafka se ejecuta en clústeres y tiene particiones y réplicas. Por lo tanto, el fallo de un solo nodo no afecta a los servicios y la capacidad de Kafka puede escalarse linealmente.

RabbitMQ también admite la implementación en clúster con múltiples opciones de cantidad de nodos.

Funciones

Tanto Kafka como RabbitMQ son middleware popular orientado a mensajes de código abierto. Se diferencian principalmente en las funciones, que se enumeran en la siguiente tabla.

Tabla 5-1 Diferencias de función entre Kafka y RabbitMQ

Función	Kafka v1.1.0/2.3.0/2.7	RabbitMQ v3.7.17
Cola de prioridad	No se admite	Se admite. Se recomienda que la prioridad se establezca en 0 - 10.
Cola retardada	No se admite	Se admite
Cola de mensajes fallidos	No se admite	Se admite
Volver a intentar	No se admite	No se admite
Modelo de recuperación	Basado en pull	Basado en pull y push
Transmisión de mensajes	Se admite	Se admite
Seguimiento de mensajes	Soporta desplazamiento y seguimiento de marca de tiempo.	No se admite. Una vez que se ha reconocido la recuperación de un mensaje, RabbitMQ será notificado de que el mensaje puede ser eliminado.
Acumulación de mensajes	Soporta un mayor rendimiento de acumulación que RabbitMQ gracias a su alto rendimiento.	Se admite
Persistencia	Se admite	Se admite
Rastreo de mensajes	No se admite	Soportado por la función firehose o el plugin rabbitmq_tracing. Sin embargo, rabbitmq_tracing reduce el rendimiento y solo debe usarse para solucionar problemas.
Filtrado de mensajes	Se admite	No se admite, pero puede ser encapsulado.
Multi-inquilinos	No se admite	Se admite
Múltiples protocolos	Solo se admite Apache Kafka.	RabbitMQ se basa en AMQP y soporta MQTT y STOMP.
Multi-idioma	Kafka está escrito en Scala y Java y soporta clientes en múltiples lenguajes de programación.	RabbitMQ está escrito en Erlang y soporta clientes en múltiples lenguajes de programación.

Función	Kafka v1.1.0/2.3.0/2.7	RabbitMQ v3.7.17
Aceleración	Soporta la limitación de los clientes productores o consumidores.	Soporta la limitación basada en el crédito de los productores, un mecanismo que activa la protección desde dentro.
Entrega de mensaje solicitada	Soporta FIFO a nivel de partición.	Soporta FIFO solo para colas de mensajes de subproceso único sin características avanzadas como colas retardadas o colas de prioridad.
Seguridad	Soporta autenticación SSL y SASL y control de permisos de lectura/escritura.	Similar a Kafka.
Idempotencia	Soporta idempotency para una sola sesión de productor.	No se admite
Mensajes de transacción	Se admite	Se admite

 **NOTA**

La comparación se hace entre Kafka de código abierto y RabbitMQ.

DMS for Kafka y DMS for RabbitMQ mantienen la compatibilidad de código abierto mientras soportan o mejoran las características en las versiones de código abierto.

6 Comparación de DMS for Kafka y Kafka de código abierto

DMS for Kafka es compatible con Kafka de código abierto y tiene características de Kafka personalizadas y mejoradas. Además de las ventajas de Kafka de código abierto, DMS for Kafka proporciona características más fiables y útiles.

Tabla 6-1 Diferencias entre DMS for Kafka y Kafka de código abierto

Categoría	Elemento	DMS for Kafka	Kafka de código abierto
Facilidad de uso	Inmediatamente disponible	Las instancias pueden crearse de forma intuitiva en cuestión de minutos y usarse inmediatamente con operaciones visualizadas y monitoreo en tiempo real.	La preparación de los recursos del servidor y la instalación y configuración del software consume mucho tiempo y es propensa a errores.
	API	Las instancias se pueden administrar fácilmente por la llamada a las RESTful API.	No disponible
Costos	Uso bajo demanda	Múltiples especificaciones están disponibles para adaptarse a diferentes necesidades. El ancho de banda de la instancia y el espacio en disco se pueden ampliar sin tiempo de inactividad.	Se incurre en gastos para configurar un servicio de mensajes y ocupar los recursos subyacentes.
	Completamente gestionado	Los servicios están disponibles sin solicitud de recursos o gastos adicionales de hardware.	Los usuarios deben preparar los recursos de hardware y configurar el servicio por sí mismos, y soportar altos costos de uso y mantenimiento.

Categoría	Elemento	DMS for Kafka	Kafka de código abierto
Éxito comprobado	Madura	DMS se ha implementado en productos de Huawei y ha demostrado su éxito en grandes eventos de comercio electrónico de Huawei, como el Vmall 11.11 Shopping Festival. También se utiliza en las nubes de los clientes de nivel de operador en todo el mundo, y cumple con los estrictos estándares de confiabilidad de nivel de operador. DMS sigue de cerca con actualizaciones de la comunidad para corregir continuamente vulnerabilidades conocidas de código abierto y agregar soporte para nuevas características.	El uso de software de código abierto requiere un largo desarrollo personal y verificación y ha tenido pocos casos de éxito.
	Amplia funcionalidad	A la vez que mantiene 100% de compatibilidad de código abierto, DMS optimiza aún más el código de código abierto para mejorar el rendimiento y la fiabilidad, y proporciona consultas de mensajes, dumping, seguimiento (disponible pronto) y muchas otras características.	La funcionalidad es limitada y requiere autodesarrollo.
Confiabilidad	Alta disponibilidad:	DMS admite la implementación entre AZ para mejorar la confiabilidad. Además, la detección automática de fallos y las alarmas garantizan un funcionamiento fiable de los servicios clave.	La alta disponibilidad requiere el autodesarrollo o la implementación de código abierto, que son costosas y no pueden garantizar la fiabilidad.

Categoría	Elemento	DMS for Kafka	Kafka de código abierto
	O&M sencillos	O&M es totalmente transparente para los inquilinos con un conjunto completo de funciones de monitoreo y alarma. El personal de O&M será informado de cualquier excepción, eliminando la necesidad de asistencia 24/7.	Los usuarios necesitan desarrollar y optimizar las funciones de operación y mantenimiento, especialmente las funciones de notificación de alarmas. De lo contrario, se requiere asistencia manual.
	Seguro	DMS utiliza aislamiento de VPC y cifrado de canal SSL.	La seguridad debe ser reforzada por los propios usuarios.

7 Notas y restricciones

DMS for Kafka tiene las siguientes restricciones, como se enumeran en [Tabla 7-1](#).

Tabla 7-1 Restricciones de uso de Kafka

Elemento	Restricciones	Descripción
Kafka Zookeeper	No expuesto externamente	DMS ZooKeeper no proporciona servicios de forma externa. Solo se utiliza para servir instancias de Kafka.
Versión	La versión del servicio puede ser 1.1.0, 2.3.0 o 2.7.	Los clientes posteriores a la versión 0.10 son compatibles. Utilice una versión que sea coherente con la versión del servicio.
Tamaño del mensaje	10 MB	El tamaño del mensaje no puede exceder los 10 MB. De lo contrario, la tarea de migración no se ejecutará.
Inicio de sesión en la VM donde residen los corredores de Kafka	No se admite	No disponible

Elemento	Restricciones	Descripción
Cantidad de particiones	Limitado	Kafka gestiona los mensajes por partición. Si hay demasiadas particiones, la creación de mensajes, el almacenamiento y la recuperación se fragmentarán, lo que afectará al rendimiento y la estabilidad. Si el número total de particiones de temas alcanza el límite superior, no podrá crear más temas.
Creación automática de temas	Se admite	Configurable durante la creación de instancias. Si está habilitado, se creará automáticamente un tema con 3 particiones y 3 réplicas cuando se crea o se recupera un mensaje de un tema que no existe.
Creación de grupos de consumidores, consumidores y productores	No se requiere	Los grupos de consumidores, los consumidores y los productores se generan automáticamente cuando se utiliza la instancia.
Disminución de la cantidad de partición	No se admite	La cantidad de partición no se puede reducir debido a las limitaciones de Apache Kafka.

8 Servicios relacionados

- **Cloud Trace Service (CTS)**
Cloud Trace Service (CTS) genera trazas para proporcionarle un historial de operaciones realizadas en recursos de servicios en la nube. Las trazas incluyen las solicitudes de operación enviadas mediante la consola de gestión o las API abiertas, así como los resultados de la operación. Puede ver todas las trazas generadas para consultar, auditar y realizar operaciones realizadas.
Para detalles sobre las operaciones registradas por CTS, consulte [Operaciones que pueden ser registradas por CTS](#)
- **VPC**
Las instancias premium de Kafka se ejecutan en VPC y utilizan las direcciones IP y el ancho de banda de la VPC. Los grupos de seguridad de las VPC mejoran la seguridad del acceso a la red a las instancias premium de Kafka.
- **Cloud Eye**
Cloud Eye es una plataforma abierta que proporciona monitoreo, informes de alarma y notificación de alarma para sus recursos en tiempo real.
📖 NOTA
Los valores de todas las métricas de instancia de Kafka se notifican a Cloud Eye cada minuto.
- **Elastic IP (EIP)**
El servicio EIP proporciona direcciones IP públicas independientes y ancho de banda para el acceso a Internet.

9 Conceptos básicos

DMS for Kafka de Huawei Cloud utiliza Kafka como motor de mensajes. Este capítulo presenta las explicaciones de los conceptos básicos de Kafka.

Tema

Un tema es una categoría para mensajes. Los mensajes se crean, se recuperan y se gestionan en forma de temas.

Los temas adoptan el patrón publicarse y suscribirse. Los productores publican mensajes en temas. Uno o más consumidores se suscriben a los mensajes de los temas. Los productores y los consumidores no están directamente vinculados entre sí.

Productor

Un productor publica mensajes sobre temas. A continuación, los mensajes se entregan a otros sistemas o módulos para su procesamiento según lo acordado.

Consumidor

Un consumidor se suscribe a mensajes en temas y procesa los mensajes. Por ejemplo, una plataforma de monitorización y alarma (un consumidor) que se suscribe para registrar mensajes en ciertos temas puede identificar registros de alarma y luego enviar notificaciones de alarma por SMS o correo electrónico.

Corredor

Un corredor es un proceso de Kafka en un clúster de Kafka. Cada proceso se ejecuta en un servidor, por lo que una delegación incluye el almacenamiento, el ancho de banda y otros recursos del servidor.

partición

Los mensajes de un tema se distribuyen en varias particiones para lograr la escalabilidad y la tolerancia a fallos.

Réplica

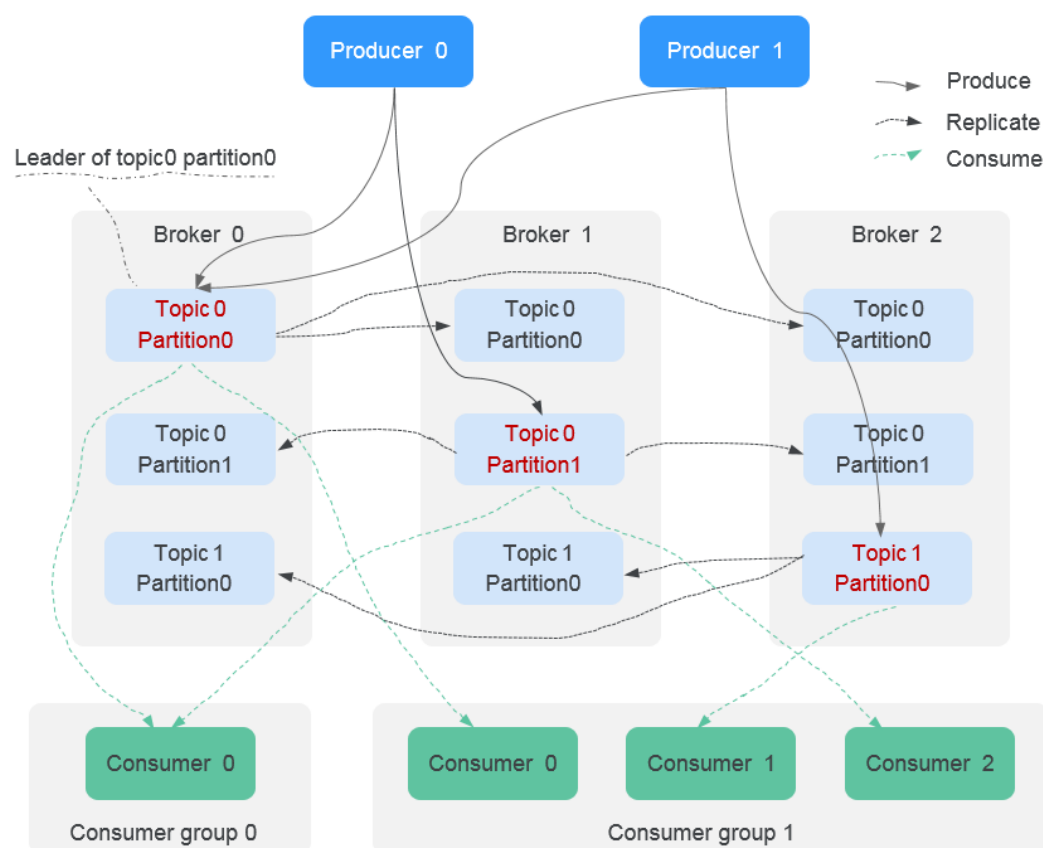
Una réplica es una copia redundante de una partición en un tema. Cada partición puede tener una o más réplicas, lo que permite la fiabilidad del mensaje.

Los mensajes en cada partición se replican y sincronizan completamente, lo que evita la pérdida de datos si una réplica falla.

Cada partición tiene una réplica como líder que maneja la creación y recuperación de todos los mensajes. El resto de réplicas son seguidores que replican al líder.

Los temas y las particiones son conceptos lógicos, mientras que las réplicas y los corredores son conceptos físicos. El siguiente diagrama muestra las relaciones entre particiones, agentes y temas en la transmisión de mensajes.

Figura 9-1 Kafka streaming de mensajes



Tiempo de envejecimiento

El período durante el que se conservan los mensajes. Los consumidores deben recuperar los mensajes antes de que finalice este período. De lo contrario, los mensajes se eliminarán y ya no se podrán recuperar.

10 Gestión de permisos

Puede utilizar Identity and Access Management (IAM) para gestionar los permisos de DMS for Kafka y controlar el acceso a sus recursos. IAM proporciona autenticación de identidad, gestión de permisos y control de acceso.

Puede crear usuarios de IAM para sus empleados y asignar permisos a estos usuarios sobre la base del principio de privilegio mínimo (PoLP) para controlar su acceso a tipos de recursos específicos. Por ejemplo, puede crear usuarios de IAM para desarrolladores de software y asignar permisos específicos para permitirles usar los recursos de DMS for Kafka, pero evitar que eliminen los recursos o realicen las operaciones de alto riesgo.

Si su cuenta de Huawei Cloud no requiere usuarios individuales de IAM para la administración de permisos, omita esta sección.

IAM es gratuito. Solo se paga por los recursos usados. Para obtener más información acerca de IAM, consulte [Descripción general del servicio IAM](#).

NOTA

Las políticas de permisos de DMS for Kafka se basan en DMS. Por lo tanto, al asignar permisos, seleccione las políticas de permisos de DMS.

Permisos de DMS for Kafka

De forma predeterminada, los nuevos usuarios de IAM no tienen ningún permiso asignado. Para asignar permisos a estos nuevos usuarios, agréguelos a uno o más grupos y adjunte directivas o roles de permisos a estos grupos.

DMS for Kafka es un servicio a nivel de proyecto implementado y accedido en regiones físicas específicas. Al asignar permisos de DMS para Kafka a un grupo de usuarios, especifique los proyectos específicos de la región donde los permisos tendrán efecto. Si selecciona **All projects**, los permisos se otorgarán para todos los proyectos específicos de la región. Al acceder a DMS for Kafka, los usuarios deben cambiar a una región en la que han sido autorizados para usar este servicio.

Puede conceder permisos a los usuarios mediante roles y políticas.

- **Roles** Un tipo de mecanismo de autorización de granularidad gruesa que proporciona solo una cantidad limitada de roles de nivel de servicio. Al usar roles para conceder permisos, también debe asignar roles de dependencia. Sin embargo, los roles no son una opción ideal para la autorización detallada y el control de acceso seguro.

- **Políticas:** Tipo de mecanismo de autorización detallado que define los permisos necesarios para realizar operaciones en recursos de nube específicos bajo ciertas condiciones. Este mecanismo permite una autorización más flexible basada en políticas para un control de acceso más seguro. Por ejemplo, puede conceder a los usuarios de DMS for Kafka solo los permisos para gestionar las instancias. La mayoría de las políticas definen permisos basados en API. Para ver las acciones API admitidas por DMS for Kafka, consulte [Políticas de permisos y acciones admitidas](#).

Tabla 10-1 enumera todas las roles definidos por el sistema y políticas admitidas por DMS for Kafka.

Tabla 10-1 Roles definidos por el sistema y políticas admitidas por DMS for Kafka

Nombre de rol/ política	Descripción	Tipo	Dependencia
DMS FullAccess	Permisos de administrador para DMS. Los usuarios con estos permisos pueden realizar todas las operaciones en DMS.	Política definida por el sistema	Ninguna
DMS UserAccess	Permisos de usuario comunes para DMS, excluyendo los permisos para crear, modificar, eliminar, dumping, y escalar instancias.	Política definida por el sistema	Ninguna
DMS ReadOnlyAccess	Permisos de sólo lectura para DMS. Los usuarios a los que se han concedido estos permisos sólo pueden ver los datos DMS.	Política definida por el sistema	Ninguna
DMS Administrator	Permisos de administrador para DMS.	Rol definido por el sistema	Este rol depende de los roles Tenant Guest y VPC Administrator .

 **NOTA**

Las políticas definidas por el sistema contienen acciones de OBS. Debido al almacenamiento en caché de datos, las políticas entran en vigor cinco minutos después de que se adjuntan a un usuario, grupo de usuarios o proyecto de empresa.

En el [cuadro 2](#) se enumeran las operaciones comunes admitidas por cada política o función del sistema de DMS for Kafka. Seleccione las políticas según sea necesario.

Tabla 10-2 Operaciones comunes admitidas por cada política definida por el sistema o función de DMS para Kafka

Operación	DMS FullAccess	DMS UserAccess	DMS ReadOnlyAccess
Creación de una instancia	√	×	×
Modificación de instancias	√	×	×
Eliminación de instancias	√	×	×
Modificación de especificaciones de instancia	√	×	×
Habilitar dumping	√	×	×
Creación de una tarea de dumping	√	√	×
Creación de instancias	√	√	×
Consulta de información de instancia	√	√	√

Enlaces útiles

- [Descripción general del servicio IAM](#)
- [Creación de grupos de usuarios, usuarios y concesión de permisos de DMS for Kafka](#)
- [Políticas de permisos y acciones admitidas](#)

11 Facturación

Huawei Cloud DMS for Kafka admite pago por uso. Para obtener más información, consulte [Detalles del precio](#).

Conceptos de facturación:

Huawei Cloud DMS for Kafka se factura según las especificaciones de instancia de Kafka y el espacio de almacenamiento. Hay dos tipos de especificaciones, una para CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg y el otro para las regiones restantes.

Tabla 11-1 Facturación de DMS for Kafka

Concepto de facturación	Descripción
instancia	<ul style="list-style-type: none"> ● Instancias de Kafka en regiones distintas de la CN Este-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg se facturan según su ancho de banda. Para garantizar una ejecución estable del servicio, se recomienda elegir un ancho de banda 30% más alto que el rendimiento real esperado para el tráfico de lectura o escritura, el que sea más alto. También debe considerar otros parámetros descritos en Tabla 11-2. Por ejemplo, si el tráfico de escritura máximo es 70 MB/s y el tráfico de lectura máximo es 200 MB/s, el ancho de banda elegido debe ser al menos $200 \text{ MB/s} \times (1 + 30\%) = 260 \text{ MB/s}$. ● Instancias de Kafka en las regiones CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg se facturan según la variante de ECS y el número de corredores. Al comprar una instancia, seleccione las variantes de ECS apropiados y el número de corredores según la evaluación del servicio. Tabla 11-3 lista el rendimiento por corredor. ● Las instancias de Kafka se pueden facturar de pago por uso (por hora). ● Si habilita la función de volcado, habrá cargos adicionales. Por ejemplo: <ul style="list-style-type: none"> - Si habilita el dumping para una instancia de 100 MB/s (3 corredores), se crearán 2 corredores más con las mismas especificaciones para el dumping y deberá pagar por ellos. - Si habilita el dumping para una instancia de c6.4u8g.cluster, se crearán 2 corredores más con la misma variante para el dumping y deberá pagar por ellos.

Concepto de facturación	Descripción
Espacio de almacenamiento	<ul style="list-style-type: none"> Las colas se facturan según el espacio de almacenamiento. Para cada tipo de especificación de instancia, puede elegir el tipo de disco de E/S común, E/S alta o E/S ultra alta para satisfacer sus requisitos de servicio. Puede especificar el número de réplicas. Por ejemplo, si el tamaño de disco requerido para almacenar datos de mensajes es de 500 GB y hay tres réplicas, la capacidad del disco debe ser al menos: 500 GB x 3 = 1500 GB. Tabla 11-3 describe las opciones de espacio de almacenamiento en incrementos de 100 GB. El espacio de almacenamiento se puede facturar de pago por uso (por hora).

Tabla 11-2 Especificaciones de instancia de Kafka

Especificación (ancho de banda)	Cantidad máxima de particiones	Grupos máximos de consumidores	Conexiones máximas de cliente	Espacio de almacenamiento
100 MB/s	300	60	3000	600 GB - 90,000 GB
300 MB/s	900	300	10,000	1200 GB - 90,000 GB
600 MB/s	1800	600	20,000	2400 GB - 90,000 GB
1200 MB/s	1800	600	20,000	4800 GB - 90,000 GB

Las especificaciones enumeradas en **Tabla 11-2** están disponibles en regiones distintas de CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg.

Tabla 11-3 Especificaciones de instancia de Kafka

Variante	Corredores	TPS máximo por corredor	Particiones máximas por corredor	Máximo de grupos de consumidores por corredor	Máximo de conexiones de cliente por corredor	Espacio de almacenamiento
c6.2u4g.cluster	3 - 30	30,000	250	20	1000	600 GB de-300,000 GB
c6.4u8g.cluster	3 - 30	100,000	500	100	2000	1200 GB-600,000 GB
c6.8u16g.cluster	3 - 30	150,000	1000	150	3000	2400 GB - 900,000 GB
c6.12u24g.cluster	3 - 30	200,000	1500	200	4000	3600 GB - 900,000 GB
c6.16u32g.cluster	3 - 30	250,000	2000	200	5000	4800 GB - 900,000 GB

Las especificaciones enumeradas en [Tabla 11-3](#) están disponibles en regiones CN East-Shanghai1, CN North-Beijing4, CN South-Guangzhou, AP-Singapore, LA-Mexico City2, LA-Santiago, and AF-Johannesburg.

Cartografía entre las variantes antiguas y nuevas

[Tabla 11-4](#) compara las variantes de instancia Kafka antiguas y nuevas.

Tabla 11-4 Cartografía entre las variantes antiguas y nuevas de instancia Kafka

Variante antigua	Particiones máximas (antiguas)	TPS (Antiguo)	Variante nueva	Particiones máximas (nuevas)	TPS (Nuevo)
100 MB/s (c6.large.2 x 3)	300	100,000	c6.2u4g.cluster x 3	750	100,000
300 MB/s (c6.xlarge.2 x 3)	900	300,000	c6.4u8g.cluster x 3	1500	300,000
600 MB/s (c6.2xlarge.2 x 4)	1800	600,000	c6.8u16g.cluster x 4	4000	600,000
1200 MB/s (c6.2xlarge.2 x 8)	1800	1.2 millones	c6.8u16g.cluster x 8	8000	1.2 millones

Suponga que tiene cuatro 8 vCPU | Máquinas virtuales de 16 GB y requieren tres réplicas y entre 10,000 y 15,000 particiones. El tráfico de escritura de cada VM es de 720 MB/s, el tráfico de lectura es de 7.2 GB/s y el tráfico de replicación es de 90 MB/s x 2. La variante y el costo de su instancia de Kafka se calculan de la siguiente manera:

Tráfico total de escritura: $720 \text{ MB/s} \times 3 \times 4 = 1080 \text{ MB/s}$

Tráfico total de lectura: $7.2 \text{ GB/s} \times 4 = 3686 \text{ MB/s}$

Tráfico total de replicación: $90 \text{ MB/s} \times 2 \times 4 = 720 \text{ MB/s}$

Tráfico total: $1080 \text{ MB/s} + 3686 \text{ MB/s} + 720 \text{ MB/s} = 5486 \text{ MB/s}$

Para garantizar los servicios estables, el tráfico total debe ser solo el 80% del ancho de banda que configure para su instancia, por lo que el tráfico total real debe ser $5486 \text{ MB/s} / 80\% = 6857.5 \text{ MB/s}$.

Si la instancia utiliza un disco de E/S ultra alta, un corredor único puede soportar un máximo de 350 MB/s de tráfico. Por lo tanto, el número de corredores requeridos es de 20 ($6857.5/350$). La variante de la instancia de Kafka debe ser **c6.8u16g.cluster x 20**. Una instancia **c6.8u16g.cluster** soporta un máximo de 1000 particiones por corredor, por lo que **c6.8u16g.cluster x 20** puede cumplir con sus requisitos de particiones de 10,000 a 15,000.

El espacio de almacenamiento mínimo de un corredor único es de 800 GB, por lo tanto, el espacio de almacenamiento total es de $800 \text{ GB} \times 20 = 16,000 \text{ GB}$.

Precio total de la instancia de Kafka: $\text{¥} 48,600$ (precio de instancia) + $\text{¥} 16,000$ (precio del espacio de almacenamiento) = $\text{¥} 64,600/\text{mes}$.

Modos de facturación

Pago por uso (por hora): Este modo de facturación es flexible, lo que le permite iniciar y detener servicios en cualquier momento. Usted paga solo por la duración real del uso. La unidad de tiempo mínima es de una hora. Menos de una hora se registra como una hora.

Cambio de configuraciones

- Puede cambiar el ancho de banda de una instancia de Kafka. A continuación, se le facturará según las nuevas especificaciones inmediatamente después del cambio.
- Puede cambiar el número de corredores para una instancia de Kafka. A continuación, se le facturará según las nuevas especificaciones inmediatamente después del cambio.
- También puede cambiar el espacio de almacenamiento de las colas de Kafka. Se le facturará según el nuevo espacio de almacenamiento inmediatamente después de su aumento. El espacio de almacenamiento solo se puede aumentar y no se puede reducir. El incremento mínimo de aumento es de 100 GB.