



Alta disponibilidad de Tableau Server

Análisis cruciales a gran escala

Por: Fatima Calcuttawala, administradora de productos
Kitty Chou, administradora de productos

Índice

El análisis de autoservicio es crucial	3
Entendiendo la alta disponibilidad	3
Implementación de alta disponibilidad mínima	4
Balanceador de carga externo (ELB)	5
Entendiendo la alta disponibilidad de Tableau Server para cada proceso de servidor ...	6
Aspectos generales	6
Servicios de TSM.....	6
Servicios empresariales.....	10
Supervisión del estado del clúster	16
Integración con herramientas de supervisión de terceros	18
Recuperación del error del nodo inicial.....	19
Consideraciones arquitectónicas	19
Implementación básica de la alta disponibilidad de 3 nodos	20
Implementaciones con más de tres nodos	21
Conjunto de servicio de coordinación externa.....	22
Más allá de la alta disponibilidad	23
Acerca de Tableau	24
Recursos adicionales	24

El análisis de autoservicio es crucial

Hoy en día, el análisis de autoservicio y la toma de decisiones basada en los datos son la norma en las organizaciones líderes de todo el mundo. Los usuarios y los encargados de tomar decisiones han llegado a depender del acceso inmediato a los datos y a las herramientas de autoservicio para responder sus preguntas en tiempo real. Los ejecutivos entienden la importancia de las decisiones basadas en datos en sus empresas y confían en estos sistemas todos los días. Esta dependencia de los datos requiere un alto grado de disponibilidad para los sistemas subyacentes. Las capacidades de una plataforma deben ser más accesibles y fácilmente configurables por los equipos existentes y las herramientas empresariales.

Tableau Server es el futuro del análisis de autoservicio crucial. Permite la exploración rápida de datos de autoservicio, promueve la confianza en el contenido y los datos a través de una gobernanza robusta y es fácil de implementar, administrar y escalar en cualquier empresa. En este documento, exploraremos cómo Tableau Server ejecuta Tableau Services Manager (TSM) para ofrecer un análisis de autoservicio a gran escala con alta disponibilidad (HA).

Entendiendo la alta disponibilidad

El objetivo de los sistemas de alta disponibilidad es minimizar el tiempo de inactividad del sistema. La mayoría de los administradores de sistemas planifican tiempo de inactividad para el mantenimiento, las actualizaciones y los parches. Además, existe la posibilidad de fallas inesperadas, lo que se conoce como tiempo de inactividad no planificado. Por supuesto, los administradores deben realizar un mantenimiento planificado para las actualizaciones de hardware o software; el objetivo es minimizar el tiempo de inactividad no planificado.

Existen dos estrategias comunes para alcanzar la alta disponibilidad. La primera es eliminar puntos únicos de error, de modo que el sistema pueda ser robusto ante errores inesperados. Sabemos que los errores ocurren en la vida real y la mejor manera de protegerse contra estos es garantizar la redundancia en el sistema. El segundo es detectar cuándo hay errores y activar mecanismos de conmutación por error confiables según sea necesario. Tableau Server emplea ambas técnicas para alcanzar la alta disponibilidad.

Entendemos lo importante que es para los usuarios ver y comprender fácilmente sus datos. También somos conscientes de que siempre habrá eventos que amenacen la disponibilidad de los sistemas de inteligencia empresarial, ya sea en lo que respecta al hardware, software, redes o incluso errores humanos. Es por eso que hemos diseñado Tableau Server para brindar una alta disponibilidad inmediata y lo hemos hecho fácil de configurar e instalar. Como mínimo, los procesos de Tableau Server se reiniciarán automáticamente para mantener su sistema en funcionamiento en caso de error del componente. Una implementación de varios nodos correctamente configurada también utiliza procesos redundantes para lograr una alta disponibilidad del servidor.

Implementación de alta disponibilidad mínima

El primer paso para obtener una alta disponibilidad es crear una instalación distribuida de Tableau Server, ya que necesitarás cierta redundancia en tu clúster. Si uno de los nodos falla, querrás asegurarte de que todavía tienes al menos uno de cada uno de los servicios centrales para ejecutar Tableau Server configurado en uno de los otros nodos del clúster. Sin embargo, dos nodos no son suficientes para la alta disponibilidad en una instalación de Tableau Server. Para lograr una alta disponibilidad, debes instalar Tableau Server en al menos 3 nodos. La razón principal de esto es que Tableau Server se basa en el concepto del cuórum para determinar si el clúster se encuentra en un estado consistente para proteger contra problemas de partición de la red. (Consulta el [teorema CAP](#) para más información). Un cuórum es solamente otra forma de decir una mayoría absoluta. Si un cuórum (o la mayoría) de los nodos en el clúster está de acuerdo con el estado del sistema, podemos concluir que el sistema es consistente y, por lo tanto, no es susceptible al problema de partición de red. Con un clúster de dos nodos, si un nodo se desactiva, entonces no hay cuórum que pueda establecerse, y Tableau Server debe entrar en un modo de error. Un clúster con tres o cuatro nodos puede tolerar la pérdida de, como máximo, un nodo. Un clúster con cinco o más nodos puede tolerar la pérdida de hasta dos nodos. (Consulta el [servicio de coordinación](#) para más detalles). Puedes agregar fácilmente nodos adicionales y configurar servicios redundantes en los nodos adicionales mediante la interfaz de usuario web de TSM o la interfaz de línea de comandos (CLI).

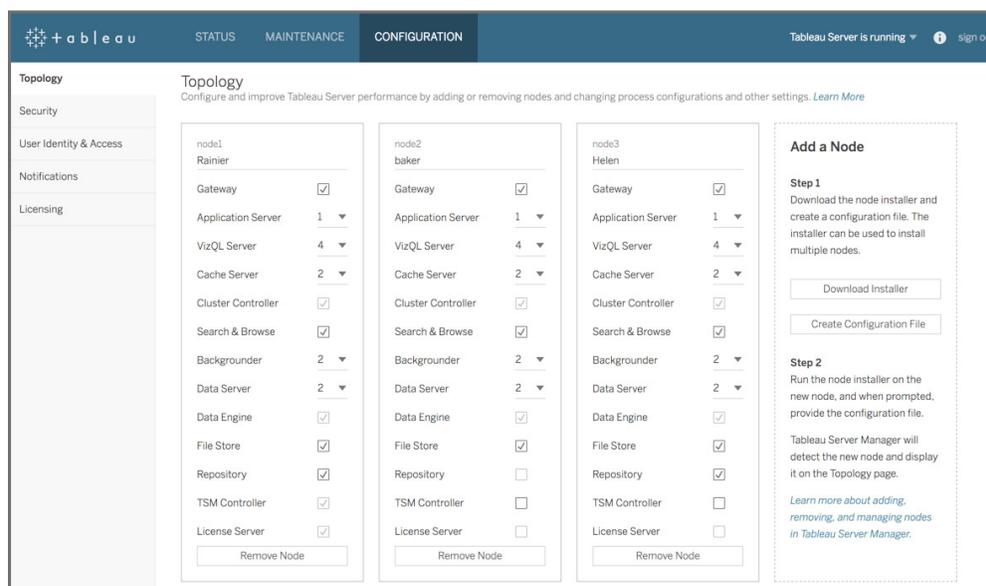


Figura 1 La interfaz de usuario web de Tableau Services Manager te permite agregar/eliminar fácilmente nodos y procesos por nodo en un clúster.

Con Tableau Server ejecutando TSM, la configuración ha cambiado de un concepto de servidor primario/de trabajo a todos los nodos como pares; es decir, todos los nodos del clúster se tratan como pares. Con TSM, no necesitas una copia de seguridad de las máquinas primarias dedicadas a la conmutación por error en caso de error en el servicio de licencias. Consulta el [servicio de licencias](#) para comprender más sobre cómo podemos lograr la alta disponibilidad.

Balancedador de carga externo (ELB)

Al implementar un sistema de alta disponibilidad, también debemos considerar cómo se conectan los usuarios finales al clúster de Tableau Server. En la mayoría de los casos, los usuarios se conectarán a través de un nombre de host DNS que proporciones a Tableau Server; este puede ser el nombre de host de cualquier nodo en el clúster con un proceso de puerta de enlace. Sin embargo, si el proceso de puerta de enlace se desactiva, los usuarios que se conecten a ese host no podrán acceder a Tableau Server. Para mejorar la confiabilidad del clúster de Tableau Server, recomendamos configurar varias puertas de enlace en el clúster y agregar un equilibrador de carga externo frente a tu clúster de Tableau Server. Los usuarios finales simplemente se conectarán al DNS para el equilibrador de carga externo, y el equilibrador de carga puede distribuir las solicitudes a través de las puertas de enlace disponibles en el clúster. Si una puerta de enlace en un clúster deja de estar disponible, el equilibrador de carga puede detectar este error y detener el envío de solicitudes. Consulta [añadir un equilibrador de carga](#) para obtener más detalles sobre cómo configurar esto correctamente para Tableau Server.

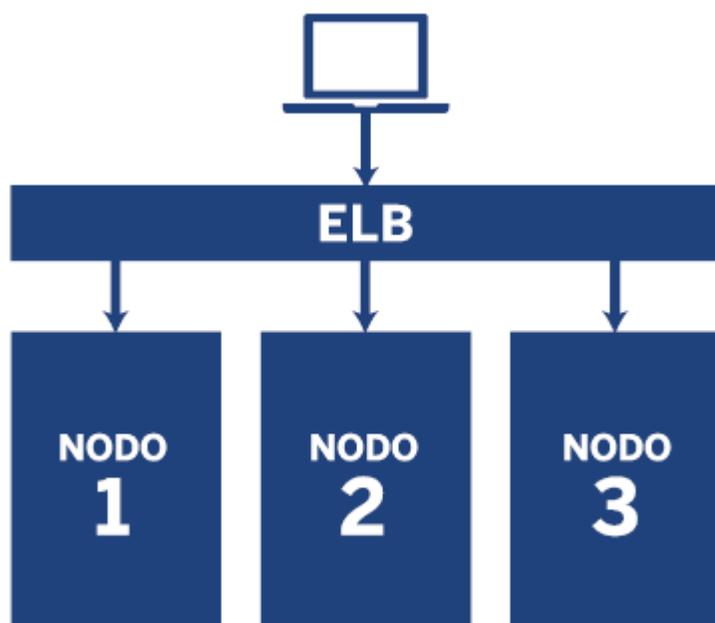


Figura 2 El clúster de Tableau con tres nodos detrás de un ELB.

Entendiendo la alta disponibilidad de Tableau Server para cada proceso de servidor

Tableau Server tiene varios procesos que garantizan que todo el sistema sea accesible para los usuarios finales. Esta sección asume que estás familiarizado con los componentes de Tableau Server y lo que hacen. Si no estás familiarizado con ellos, te sugerimos que primero leas sobre los [procesos de Tableau Server](#).

Comprender cómo hacer que Tableau Server esté altamente disponible en gran medida incluye comprender cómo hacer que cada uno de sus componentes esté altamente disponible. En otras palabras, se debe garantizar una alta disponibilidad para cada componente a fin de que todo el clúster de Tableau Server esté altamente disponible, proporcionando redundancia para puntos únicos de error. Consideremos cada componente a su vez.

Aspectos generales

Tableau Server tiene automatización incorporada para reiniciar todos sus procesos de servidor. Esta automatización garantiza una alta disponibilidad al reiniciar automáticamente los procesos del servidor con errores. El hardware o la máquina virtual que aloja Tableau Server debe estar en buen estado para que ocurra esta automatización.

Para protegerse contra el error de un nodo completo, es importante configurar los procesos únicos del servidor para que sean redundantes en los diferentes nodos del clúster. Esta redundancia es posible para todos los procesos, excepto para el servicio de licencias y el servicio del controlador de administración, que solamente puede tener una sola instancia ejecutándose en todo el clúster. Discutimos cómo abordar estas dos limitaciones de proceso con más detalle en la sección conmutación por error del nodo inicial.

Servicios de TSM

En la versión Tableau 10.5 de Linux y Tableau 2018.2 de Windows, presentamos un nuevo conjunto de servicios con Tableau Services Manager (TSM). El propósito de estos servicios es administrar la implementación de Tableau Server. Hay cinco servicios de TSM que siempre están en ejecución, incluso cuando Tableau Server se ha detenido. Otros tres servicios de mantenimiento de TSM se detienen, excepto cuando sea necesario. Si los servicios de TSM se detienen, estos se reinician automáticamente si la computadora host está en buen estado.

Servicio de coordinación

El servicio de coordinación está basado en Apache Zookeeper, un proyecto de código abierto, y coordina las actividades en el servidor, garantizando un quórum en el caso de un error y sirviendo como la fuente de "verdad" con respecto a la topología, configuración y estado del servidor. El servicio se instala automáticamente en el nodo inicial de Tableau Server, pero no se instalan instancias adicionales a medida que agregas nodos adicionales. Debido a que el funcionamiento exitoso de Tableau Server depende de un servicio de coordinación que funcione correctamente, recomendamos que para las instalaciones de servidor de tres o más nodos, agregue instancias adicionales del servicio de coordinación al implementar un nuevo conjunto de servicio

de coordinación. Esto proporciona redundancia y disponibilidad mejorada si una instancia del servicio de coordinación tiene problemas.

La cantidad de nodos recomendados para instalar el servicio de coordinación depende de la cantidad de nodos en el clúster, como se describe en la siguiente tabla:

Nº. de nodos en el clúster	Nº. de nodos con servicio de coordinación
1-2	1
3-4	3
5+	5

Puedes seleccionar en qué nodos implementar el conjunto del servicio de coordinación, y puedes implementar un conjunto del servicio de coordinación de 1, 3 o 5 nodos.

Se producirá una interrupción completa de Tableau Server si el número de procesos en ejecución del servicio de coordinación no constituye un cuórum, que se basa en el número total de procesos de servicios de coordinación configurados. En particular, un clúster con solamente dos nodos no puede tolerar la pérdida de un proceso de servicio de coordinación individual, y esta es la razón por la que la alta disponibilidad total, incluida la conmutación por error automática, requiere un mínimo de tres nodos.

¿Qué sucede cuando falla un proceso de servicio de coordinación? Nada, siempre que el número de procesos de servicio de coordinación restantes siga constituyendo un cuórum. Si la cantidad de servicios de coordinación en funcionamiento sigue siendo inferior a un cuórum, la totalidad del clúster de Tableau Server deja de estar disponible para proteger la integridad referencial de la base de datos de Postgres y el estado del servicio subyacentes. Los procesos del servicio de coordinación con errores se reinician automáticamente si la computadora está en un estado correcto de funcionamiento.

Agente

El agente de administración supervisa el servicio de coordinación en busca de cambios en la configuración o topología y entrega nuevas configuraciones a cada servicio (configuración) o implementa nuevos servicios y elimina los antiguos (topología). El agente también verifica el estado de cada uno de los servicios y lo reporta al servicio de coordinación. Este proceso se configurará automáticamente para cada nodo del clúster durante la instalación; no se requiere una configuración explícita.

¿Qué sucede cuando un proceso de agente falla? Todos los demás procesos de Tableau Server que se ejecutan en el mismo nodo se mostrarán como "no disponibles" en la página de estado de TSM. Tableau Server continuará funcionando como se esperaba; sin embargo, no podrás realizar cambios de configuración/topología en el clúster. Los procesos con errores del agente se reinician automáticamente siempre que la computadora en sí funcione adecuadamente. Si el agente no se inicia en el nodo, puedes intentar iniciar los servicios manualmente ejecutando los siguientes comandos:

- Windows: `sc start tabadminagent_0`
- Linux: `systemctl start tabadminagent_0`

Administrador de servicios

El administrador de servicios administra el ciclo de vida de los procesos empresariales de Tableau Server que se describen en la siguiente sección. Como agente, este proceso se configurará automáticamente para usted en cada nodo del clúster durante la instalación, no se requiere una configuración explícita

¿Qué sucede cuando falla un proceso de administrador de servicios? Todos los servicios empresariales en ese nodo también fallarán. Dependiendo de los servicios que se ejecutan en el nodo, Tableau Server puede ejecutarse en un estado degradado. El administrador de servicios se reiniciará automáticamente si la máquina está en buen estado. Si el administrador de servicios no se inicia en el nodo, puedes intentar iniciar los servicios manualmente ejecutando los siguientes comandos:

- Windows: `sc start tabsvc_0`
- Linux: `systemctl start tabsvc_0`

Servicio de archivo de cliente (CFS)

El servicio de archivo de cliente almacena y distribuye los archivos que necesita TSM (es decir, certificados SSL, archivos de personalización, etc.). Funciona de manera muy similar a como lo hace el almacén de archivos para los archivos necesitados por los servicios empresariales (es decir, extracciones, miniaturas, etc.). De forma predeterminada, el CFS solamente se instala en el nodo inicial de su instalación de Tableau Server. Para configurar el CFS para que sea de alta disponibilidad, le recomendamos que configure una instancia de CFS en cada uno de los nodos donde implementa el servicio de coordinación.

¿Qué sucede cuando falla un proceso CFS? Nada, siempre y cuando todavía haya al menos un proceso de CFS en funcionamiento en el clúster. El controlador redirigirá las solicitudes de transferencia de archivos al otro proceso de CFS en funcionamiento.

Los procesos con errores del servicio de archivo de cliente se reinician automáticamente siempre que la computadora en sí funcione adecuadamente.

Servicios de mantenimiento del TSM

Existen tres servicios de mantenimiento del TSM que se instalan en cada nodo del clúster: Mantenimiento de bases de datos, copia de seguridad/restauración e importación/exportación de sitio. Estos servicios permanecen detenidos a menos que el administrador del servidor inicie una tarea de mantenimiento específica que los requiera. No se requiere una configuración de alta disponibilidad adicional para estos servicios. Estos servicios se usan solamente para tareas de mantenimiento como la copia de seguridad y restauración, y no deben afectar el funcionamiento de Tableau Server para los usuarios finales.

Controlador

El proceso del controlador de administración aloja la API REST de TSM para configurar y administrar su implementación de Tableau Server. Solamente puede haber una instancia única del controlador de administración en todo el clúster. Si el controlador de administración falla, el clúster de Tableau Server debe continuar funcionando; sin embargo, no podrá realizar cambios ni actualizaciones en la configuración/topología hasta que el controlador vuelva a funcionar. Al igual que otros servicios del TSM, el controlador se reinicia automáticamente si se detiene/falla.

Si esto no resuelve la situación, debe mover el proceso del controlador a otro nodo del clúster para mitigar la situación. Encuentra más información en la sección Recuperación del error del nodo inicial.

Servicio de licencias

El servicio de licencias es responsable de garantizar que el servidor en general tenga la licencia adecuada. Al igual que en el controlador, solamente puede haber una instancia del servicio de licencias ejecutándose en todo el clúster. El servicio de licencias debe ubicarse junto con el controlador. Si el servicio de licencias falla, Tableau Server continuará ejecutándose hasta un máximo de 72 horas hasta la próxima verificación de licencias programada o hasta que se reinicie un servicio con licencia. En caso de que falle el servicio de licencias, debes mover el servicio de licencias a otro nodo y seguir los pasos para reactivar tu clave de licencias para evitar entrar en un estado sin licencia. Encuentra más información en la sección Recuperación del error del nodo inicial.

Servicios empresariales

Puerta de enlace

El proceso de puerta de enlace se puede ejecutar en cualquiera y en todos los nodos de Tableau Server. La puerta de enlace es un servidor web que maneja todas las solicitudes a Tableau Server desde los navegadores, Tableau Desktop y otros clientes. Cada nodo solamente puede tener un proceso de puerta de enlace individual, por lo que recomendamos configurar una instancia de la puerta de enlace en varios nodos en el clúster para redundancia en el sistema. De hecho, te recomendamos que configures un proceso de puerta de enlace en cada nodo. Esto mitiga el riesgo de que el proceso de puerta de enlace sea un punto único de error que lleve a una falta de disponibilidad del servicio. Siempre que haya al menos un proceso de puerta de enlace que se ejecute y reciba tráfico, el clúster puede continuar atendiendo las solicitudes de los usuarios.

La puerta de enlace también cumple una función importante al proporcionar la alta disponibilidad para Tableau Server, ya que es el componente responsable del equilibrio de la carga o la redirección del tráfico a servicios empresariales específicos con Tableau Server. La carga de la puerta de enlace equilibra el tráfico a través del conjunto de servicios activos del mismo tipo. Detectará cuándo ciertos procesos no están disponibles y dejará de enviar tráfico de usuarios a esos procesos no disponibles durante algún tiempo. La puerta de enlace revisará periódicamente el estado de esos procesos y los agregará nuevamente en rotación una vez que se hayan recuperado. De esta manera, el tráfico se carga automáticamente de manera equilibrada y se enruta solamente a servicios en buen estado en el clúster sin ninguna intervención manual.

Entonces, ¿qué sucede cuando falla un proceso de puerta de enlace? Como se mencionó anteriormente, si ningún proceso de puerta de enlace se está ejecutando, todo el clúster de Tableau Server dejará de estar disponible. Si otros procesos de puerta de enlace siguen ejecutándose, las solicitudes que se realicen a las puertas de enlace que funcionan se procesarán normalmente. Sin embargo, cualquier solicitud recibida por la puerta de enlace en error no será redirigida y continuará produciendo un error, a pesar de la presencia de otras puertas de enlace que funcionan. Los procesos de puerta de enlace con error se reinician automáticamente; por lo tanto, mientras la computadora en sí funcione, el proceso de puerta de enlace con error se reiniciará y se reanudarán las solicitudes de servicio.

Para que tu instalación sea aún más robusta ante los errores de puerta de enlace, te recomendamos que coloques tu clúster de Tableau Server detrás de un equilibrador de carga externo para garantizar que las solicitudes solamente se enruten a los procesos de puerta de enlace que funcionan. El artículo [Añadir un equilibrador de carga](#) proporciona una guía adicional para configurar esto.

Repositorio y controlador de clúster

El repositorio es una base de datos PostgreSQL que es el repositorio central de metadatos para Tableau Server que incluye información del usuario, permisos, libros de trabajo, fuentes de datos, programaciones, etc. Si no hay un repositorio en pleno funcionamiento, el clúster completo de Tableau Server no estará disponible. En cualquier momento, debería haber un repositorio "activo", que debería estar en pleno funcionamiento. Todas las operaciones relacionadas con el repositorio utilizan este repositorio activo.

Para mejorar la disponibilidad, puedes configurar Tableau Server con un repositorio adicional, "pasivo", en un nodo diferente del clúster. Los contenidos del repositorio activo se transmiten constantemente al repositorio pasivo. En el caso de una falla del repositorio activo, un clúster configurado para la alta disponibilidad promoverá automáticamente el repositorio pasivo al estado activo para la disponibilidad continua del servidor. Los clientes que desean alta disponibilidad seguramente querrán configurar un repositorio pasivo. En un clúster completo, puede haber un máximo de solamente dos repositorios (uno activo y uno pasivo) no pueden estar en el mismo nodo y, por lo tanto, deben estar en nodos separados.

El controlador de clúster es un componente de servidor requerido para cualquier nodo en el clúster. Este proceso se configurará automáticamente para cada nodo del clúster durante la instalación; no se requiere una configuración explícita. La función principal del controlador de clúster es supervisar y administrar la puesta en marcha, el apagado y la conmutación por error de activo a pasivo del repositorio. Si un proceso del controlador de clúster presenta un error, cualquier proceso del repositorio que se ejecute en ese nodo tampoco estará disponible. Al igual que otros procesos del servidor, un proceso de controlador de clúster con errores se reiniciará automáticamente siempre que la computadora funcione adecuadamente. Cuando el controlador de clúster se reinicia, también reinicia cualquier proceso del repositorio configurado en ese nodo. Por ello, un error en el proceso del controlador de clúster puede causar un problema con el repositorio.

¿Qué sucede cuando se produce un error en un proceso de repositorio? Eso depende.

La siguiente lista resume los diversos casos:

- Si se produce un error en el repositorio pasivo, los usuarios no deberían experimentar ningún impacto. Todo seguirá funcionando ya que el repositorio activo todavía está funcionando. En segundo plano, el repositorio pasivo se reiniciará y la replicación de datos se reanudará, aunque puede haber algún retraso antes de que el repositorio pasivo vuelva a estar completamente sincronizado con el repositorio activo.
- Si se produce un error en el repositorio activo, y hay un repositorio pasivo completamente sincronizado disponible, y el clúster está configurado para alta disponibilidad, se activará automáticamente una conmutación por error en el repositorio pasivo después de cinco minutos. Después de la conmutación por error, el repositorio pasivo anterior será el nuevo repositorio activo. El sistema reiniciará el repositorio activo anteriormente fallido como el nuevo repositorio pasivo y comenzará la sincronización. También reiniciará otros procesos relevantes automáticamente, de modo que estén al tanto del repositorio activo recientemente promovido y puedan volver a conectarse. Durante esta breve ventana de reinicios, los usuarios experimentarán una interrupción del servicio. Sin embargo, los administradores no tienen que intervenir para garantizar la disponibilidad continua ya que esta es una secuencia automatizada. Si deseas promover manualmente el repositorio de pasivo a activo, puedes hacerlo usando el siguiente comando:

```
tsm topology failover-repository
```

- Si se produce un error en el repositorio activo, y si no hay un repositorio pasivo completamente sincronizado, Tableau Server no estará disponible hasta que se pueda reiniciar el repositorio activo. El sistema intentará hacer esto automáticamente, pero dependiendo de la razón del error, esto podría no ser posible. Esto significa que si solamente se puede usar actualmente el repositorio activo, entonces Tableau Server no puede considerarse de alta disponibilidad. El repositorio activo es un punto único de error para todo el sistema si no hay un repositorio pasivo sincronizado.

Ten en cuenta lo siguiente: PostgreSQL realiza y controla la sincronización entre dos repositorios.

Tableau Server no tiene control sobre esto.

Almacén de archivos

El proceso del almacén de archivos gestiona el almacenamiento y la replicación de extracciones y otros archivos entre nodos. Para que el almacén de archivos sea altamente disponible, simplemente configura el proceso del almacén de archivos en varios nodos del clúster.

Cómo funciona el almacén de archivos

Se crea un archivo de extracción en el sistema cuando un usuario lo publica por primera vez en Tableau Server o se produce una actualización de extracción. Inmediatamente después de uno de estos eventos, la extracción existe en un único almacén de archivos. No se puede decir que el archivo de extracción específico esté altamente disponible todavía, ya que no tiene redundancia y, por lo tanto, es un punto único de error. Los procesos del almacén de archivos se comunican entre sí para replicar rápidamente las extracciones locales a todos los demás nodos del almacén de archivos en el clúster. El proceso del almacén de archivos está diseñado para copiar los archivos tan rápido como lo permitan los recursos de la red del clúster, pero puede tomar una cantidad variable de tiempo dependiendo del tamaño de la extracción y otras demandas de recursos en el sistema. Una vez que la copia está disponible en varios nodos dentro de los clústeres, el archivo de extracción se replica y tiene redundancia.

¿Qué sucede si se produce un error en un proceso de almacén de archivos? Hay dos consecuencias:

- La copia de los archivos de extracción hacia y desde el nodo afectado se detiene.
- La eliminación de los archivos de extracción que ya no son necesarios en el nodo afectado se suspende. (Este proceso de eliminación generalmente se conoce como "recojo de la extracción").

La suspensión del recojo de la extracción no tiene un efecto inmediato. Simplemente significa que una acumulación de archivos de extracción no deseados comenzará a consumir espacio en disco en ese nodo. Eventualmente, esto puede causar un problema, pero los nodos del tamaño correcto deberían tener un cómodo búfer de espacio en disco.

La falta de replicación de archivos significa que los nuevos archivos de extracción agregados al nodo del almacén de archivos de trabajo no estarán disponibles en los nodos del almacén de archivos con errores en el clúster. Una vez que se reinicia el proceso del almacén de archivos, el sistema se corrige a sí mismo al garantizar la sincronización entre todos los almacenes de archivos en todos los nodos.

Los procesos con errores del almacén de archivos se reinician automáticamente siempre que la computadora en sí funcione adecuadamente. El proceso del almacén de archivos debe reanudar rápidamente su función, incluida la sincronización de todos los archivos, ya sea que se hayan agregado durante el período del error o después.

Servidor de aplicaciones

El servidor de aplicaciones (VizPortal) maneja la aplicación web, las llamadas a la API REST y admite la navegación y la búsqueda. Lograr una alta disponibilidad con el servidor de aplicaciones es fácil. Solamente necesitas configurar las instancias del servidor de aplicaciones en cada nodo en el clúster de Tableau Server.

¿Qué sucede si se produce un error en un proceso de servidor de aplicaciones? Las solicitudes gestionadas por esa instancia fallarán, pero las solicitudes posteriores se enviarán a otros procesos de servidor de aplicaciones en ejecución. Suponiendo que el nodo que contiene el servidor de aplicaciones con errores aún se está ejecutando, el proceso fallido debería reiniciarse automáticamente en segundos.

Servicio SAML

Para las instalaciones de Tableau Server que tienen habilitado **SAML de un sitio específico** también se ejecutará una instancia del servicio SAML en cada nodo configurado con el servidor de aplicaciones. Esto se configurará automáticamente cuando se haya habilitado SAML de un sitio específico en el servidor. El servicio SAML en Tableau Server se mostrará como detenido a menos que el sitio SAML esté habilitado. Si este proceso entra en un estado de error, los usuarios no podrán iniciar sesión en Tableau Server si la solicitud del usuario se enruta a un servidor de aplicaciones en este nodo. Al igual que otros procesos, cuando se produce un error en un nodo que contiene el servicio SAML, el proceso fallido debería reiniciarse automáticamente en segundos.

Procesador en segundo plano

El procesador en segundo plano ejecuta las tareas del servidor, incluidas las actualizaciones de la extracción, las suscripciones, las tareas “Ejecutar ahora” y las tareas iniciadas desde tabcmd. Para que el servicio del procesador en segundo plano esté altamente disponible, debes configurar una o más instancias para que se ejecuten en varios nodos del clúster. Cuando decidas dónde y cuántos procesadores en segundo plano se ejecutarán, considera cómo los otros procesos de servidor afectarán a la capacidad disponible de cada máquina.

¿Qué pasaría si se produce un error en un proceso de procesador en segundo plano? Los trabajos de actualización y de suscripción en el proceso de procesador en segundo plano con errores se reintentan una vez que el proceso de procesador en segundo plano se recupera del error. La mayoría de los trabajos en segundo plano están programados para ejecutarse periódicamente, y la misma tarea en segundo plano será recogida y ejecutada normalmente en la próxima hora programada por un proceso de procesador en segundo plano que funcione.

Los procesos de procesador en segundo plano con errores se reinician automáticamente siempre que la computadora esté funcionando bien y los trabajos con errores se reintentarán.

Servidor de datos

El servidor de datos administra las conexiones a las fuentes de datos de Tableau Server. Para que el servidor de datos esté altamente disponible, configure uno o más procesos del servidor de datos para que se ejecuten en varios nodos del clúster.

¿Qué sucede si se produce un error en un proceso de servidor de datos? Las consultas que se ejecutan a través del proceso del servidor de datos fallarán, lo que resultará en una vista, una actualización de extracción o una alerta con errores. Las solicitudes subsiguientes, incluido el reintento de la operación con errores, deberían tener éxito siempre que haya un servidor de datos en funcionamiento disponible para aceptar solicitudes redirigidas.

Tableau Server no depende del servidor de datos para funcionar; sin embargo, sin un servidor de datos en ejecución, los libros de trabajo en el servidor pierden la capacidad de conectarse o consultar las fuentes de datos publicadas. Cualquier vista que no utilice el servidor de datos para cualquiera de sus fuentes de datos debería funcionar correctamente.

VizQL Server

El VizQL Server carga y representa vistas, y calcula y ejecuta consultas. Para lograr una alta disponibilidad para el proceso del VizQL Server, simplemente configura una o más instancias para que se ejecuten en varios nodos.

¿Qué pasa si se produce un error en un proceso de VizQL Server? Si solamente hay un proceso de VizQL Server y este produce un error, Tableau Server ya no podrá representar ninguna vista. La alta disponibilidad requiere configurar procesos VizQL redundantes. Una configuración bastante típica consiste en dos a cuatro procesos de VizQL Server en cada nodo. Esto al mismo tiempo satisface la necesidad de alta disponibilidad y de escalabilidad. Si se están ejecutando varios procesos de VizQL Server, el error en un único proceso resultará en la falla de cualquier solicitud y la pérdida de datos de la sesión en el momento del error. Cualquier solicitud futura se dirigirá a los otros procesos de VizQL Server en funcionamiento en el clúster de Tableau Server.

Motor de datos

El motor de datos procesa las cargas y consulta las extracciones de datos cuando se utilizan análisis en memoria. Un proceso de motor de datos se configurará automáticamente en cualquier nodo que ejecute uno o más procesos de almacén de archivos, VizQL Server, procesador en segundo plano, servidor de datos o servidor de aplicaciones.

¿Qué sucede cuando se produce un error en un proceso del motor de datos? Se producirá un error en las actualizaciones de la extracción. Las nuevas actualizaciones de extractos de .tde no se verán afectadas; sin embargo, las actualizaciones de la extracción de .hyper en ese nodo continuarán fallando hasta que se recupere el motor de datos. Si la representación de la visualización está ejecutando cualquier consulta en el nodo con errores, la representación fallará. Ejecutar de nuevo la misma operación hará que se reasigne automáticamente a un motor de datos diferente que sí esté funcionando. Si la visualización incluye una extracción de respaldo y la solicitud se enruta a un nodo con un motor de datos con errores, entonces la representación de la vista continuará fallando hasta que se recupere el motor de datos en ese nodo.

Los procesos de motor de datos con errores se reinician automáticamente; siempre que la computadora esté funcionando bien, el proceso del motor de datos se reiniciará.

Servidor de caché

El servidor de caché proporciona un caché de consulta externo compartido. Es un caché de pares de clave/valor que contienen información de consultas anteriores para acelerar las solicitudes futuras. Para que el servidor de caché esté altamente disponible, configure uno o más procesos del servidor de caché en varios nodos del clúster.

¿Qué sucede cuando el proceso del servidor de caché falla? Las consecuencias son relativamente leves. Tableau Server aún funcionará, pero las acciones pueden tardar más tiempo, ya que no tienen resultados disponibles de caché previo. A medida que las consultas se vuelven a ejecutar, el servidor de caché reiniciado se vuelve a llenar, lo que eventualmente acelera las cosas para los usuarios finales. En efecto, el servidor de caché no tiene un impacto en la disponibilidad; sin embargo, tiene un impacto en varios escenarios de rendimiento del usuario final. Para reducir el impacto en el rendimiento del usuario, ejecute varios procesos de este tipo en todo el clúster.

Un proceso con error del servidor de caché se reinicia automáticamente; siempre que la computadora esté en buen estado, el proceso del servidor de caché se reiniciará.

Buscar y examinar

El servicio de búsqueda maneja la búsqueda rápida, el filtro, la recuperación y la visualización de metadatos de contenido en el servidor. Lograr una alta disponibilidad para el proceso de búsqueda y exploración es fácil. El sistema debe configurarse de modo que el proceso de búsqueda y exploración se ejecute en varios nodos.

¿Qué sucede si falla un proceso de búsqueda y exploración? Tableau Server estará en gran parte utilizable, y aunque los usuarios todavía pueden iniciar sesión en el sistema, el contenido del libro de trabajo parecerá estar faltante. El contenido realmente no está faltante, simplemente no se está devolviendo en los resultados de búsqueda; se mostrará de nuevo después de que se reinicie el proceso de búsqueda y exploración. Si se configura más de un proceso de búsqueda y exploración y se ejecuta en varios nodos cuando se produce el error, las solicitudes realizadas a un proceso de búsqueda y exploración fallido también fallarán, pero las solicitudes subsiguientes se enviarán a procesos de búsqueda y exploración que funcionen. Cada proceso de búsqueda y exploración se indexa en todos los nodos del clúster, por lo tanto, si todos los procesos de búsqueda y exploración menos uno fallan, los resultados se devolverán en todos los nodos.

Supervisión del estado del clúster

Hasta ahora, hemos visto cómo se comporta cada proceso de servidor en caso de error y cómo puedes mitigar los riesgos de falla para garantizar que el clúster de Tableau Server esté altamente disponible. Además de planificar cada escenario de error, también debes supervisar proactivamente el clúster para detectar los errores que hayan ocurrido en el pasado.

Los administradores del sistema pueden supervisar el estado de Tableau Server en la página de estado de TSM. Esto proporciona a los administradores una comprensión de la salud del clúster general al mostrar los procesos de servidor en todos los nodos.

Process	node1	node2	node3
Gateway	✓	✓	✓
Application Server	✓	✓	✓
VizQL Server	✓ ✓ ✓ ✓ ✓	✓ ✓ ✓ ✓ ✓	✓ ✓ ✓ ✓ ✓
Cache Server	✓ ✓	✓ ✓	✓ ✓
Cluster Controller	✓	✓	✓
Search & Browse	✓	✓	✓
Backgrounder	✓ ✓	✓ ✓	✓ ✓
Data Server	✓ ✓	✓ ✓	✓ ✓
Data Engine	✓	✓	✓
File Store	✓	✓	✓
Repository	✓		✓
TSM Controller	✓		
License Server	✓		

Refresh Status ✓ Active ⌛ Busy ❌ Error ✕ Stopped

Figura 3 La página de estado de TSM muestra el estado de cada proceso en cada nodo.

Tableau Server puede configurarse para enviar alertas de correo electrónico a los administradores de servidor cuando se produce un error en el sistema. Esto también se puede utilizar para proporcionar una advertencia temprana de problemas de capacidad de espacio en disco.

Notifications
Configure Tableau Server to send email notifications about critical events, processes and server health. Email notifications must be sent through an email (SMTP) server. [Learn more](#)

Email Server **Events**

Events
You can specify which server events will trigger an email notification. We recommend enabling all notifications. [Learn more](#).

Content updates
 Send email when flow runs, encryption jobs, or scheduled refreshes fail
 Allow users to receive email for views that they have subscribed to

Server health monitoring
 Send emails for Tableau Server process events (up, down, and fallover)
 Send emails for Tableau Server license reporting

Drive space
 Send emails when unused drive space drops below thresholds
 Warning threshold: %
 Critical threshold: %
 Send threshold alert every: minutes
 Record disk space usage information and threshold violations for use in custom administrative views

Cancel Save Pending Changes

Figura 4 La interfaz de usuario web de TSM proporciona la capacidad de configurar alertas cuando se producen errores del sistema.

Integración con herramientas de supervisión de terceros

Además de supervisar el estado del sistema mediante los mecanismos integrados de Tableau Server, también puedes recibir de forma remota una versión en formato JSON legible por la máquina del estado de cada proceso. Deberás acceder a la API REST de TSM para cada estado. Aquí hay un ejemplo:

```
GET /status
Response status code
200 OK
Response body example
{
  "clusterStatus": {
    "nodes": [
      {
        "services": [
          {
            "serviceName": "filestore",
            "instances": [
              {
                "code": "ACTIVE",
                "processStatus": "Active",
                "instanceId": "0",
                "timestampUtc": 1497060680268,
                "currentDeploymentState": "Enabled",
                "binaryVersion": "<build>"
              }
            ],
            "rollupStatus": "Running",
            "rollupRequestedDeploymentState": "Enabled"
          }, {
            "..."
          }
        ],
        "nodeId": "node1",
        "rollupStatus": "Running",
        "rollupRequestedDeploymentState": "Enabled"
      }
    ],
    "href": "/api/0.5/status",
    "rollupStatus": "Running",
    "rollupRequestedDeploymentState": "Enabled"
  }
}
```

Recuperación del error del nodo inicial

Al instalar Tableau Server, los procesos del servicio de licencias y del controlador se instalarán de forma predeterminada en el nodo inicial del clúster. Como solamente puedes tener una instancia de cada uno de estos dos procesos en todo el clúster, no estarán disponibles en ningún otro nodo del clúster y, por lo tanto, se convertirán en puntos únicos de error en el sistema. Cuando estos procesos fallan, el servidor podría quedar sin licencia y deshabilitado. Además, no podrás administrar ni realizar cambios en el clúster de servidores hasta que te recuperes. Por lo tanto, es crucial usar las técnicas de supervisión anteriores para detectar y reaccionar ante tales errores.

Cuando detectas un error en el nodo inicial (o en un nodo que ejecuta las licencias y el controlador), Tableau Server continuará funcionando hasta la próxima verificación de licencia o hasta que se reinicien los servicios con licencia. Debes intentar identificar qué causó el problema y corregir cualquier nodo o servicio con problemas. Si no puedes llevar el nodo o los servicios a un estado normal, te recomendamos que sigas los pasos documentados en el tema [Recuperación a partir de un error de un nodo inicial](#) para mover estos dos servicios a uno de los otros nodos de tu clúster. Además, estos pasos pueden automatizarse y programarse para facilitar la repetición y evitar errores manuales durante la ejecución de estos.

Consideraciones arquitectónicas

Para una implementación exitosa de alta disponibilidad, debes comprender los objetivos de tiempo de actividad y los niveles de servicio esperados de tu organización, y luego planificar la redundancia en consecuencia. La determinación de la configuración óptima de tu clúster dependerá de las necesidades de tu empresa y de los recursos disponibles en tu organización. Si bien existen varias configuraciones posibles de alta disponibilidad para el éxito, es importante comprender los intercambios entre grupos de diversos tamaños para que puedas tomar la mejor decisión para tu entorno.

Además, puedes elegir entre configuraciones de computadora simétricas (cada nodo tiene todos los mismos componentes) o asimétricas (cada nodo tiene un número y tipo de componentes variables). Muchas veces, tener una configuración simétrica para los nodos de trabajo simplifica la clonación y la adición de nodos adicionales al clúster. Sin embargo, debes tener en cuenta las restricciones de un repositorio individual activo e individual pasivo mientras planificas tu configuración.

Implementación básica de la alta disponibilidad de 3 nodos

Como hemos comentado anteriormente, debes tener al menos tres nodos en el clúster para que se ejecuten completamente en modo de alta disponibilidad. Un clúster de 3 nodos también es un excelente punto de partida para implementaciones más grandes. Si un nodo falla, aún tendrá un cuórum con los dos nodos activos restantes. La Figura 5 a continuación muestra una implementación básica de alta disponibilidad de 3 nodos. Como se discutió, el nodo inicial (Nodo 1) aloja las funciones administrativas y de licencia para todo el clúster, pero aparte de eso, no hay nada especial en el nodo inicial frente a los otros nodos del clúster. En esta configuración, si el Nodo 1 con el repositorio activo falla, el Nodo 3 con el repositorio pasivo se activa automáticamente. Los tres nodos del clúster están configurados para alimentar los datos y la funcionalidad de visualización de Tableau Server con redundancia en todo el clúster y dentro de un nodo individual. El servicio de coordinación se implementa en los tres nodos. Si alguno de los nodos falla, todavía habría cuórum (dos de cada tres nodos siguen funcionando). La combinación exacta de estos componentes puede variar según tus necesidades de escalabilidad además de tus necesidades de alta disponibilidad. Para obtener más información sobre la escalabilidad de Tableau Server, también puedes consultar nuestro [informe](#).



Figura 5 Ejemplo de una implementación básica de alta disponibilidad de 3 nodos.

Además, debes implementar el clúster detrás del equilibrador de carga externo para mejorar la disponibilidad del usuario final. Si recién estás empezando con la alta disponibilidad, ésta es una buena configuración para comenzar.

Implementaciones con más de tres nodos

Al implementar un clúster con un número par de nodos (por ejemplo, cuatro o seis nodos), la capacidad de cuórum de este clúster será la misma que la de un clúster con un nodo menos. Los servicios de coordinación solamente se pueden implementar en números impares de nodos. Consideremos un ejemplo de un clúster de 4 nodos.

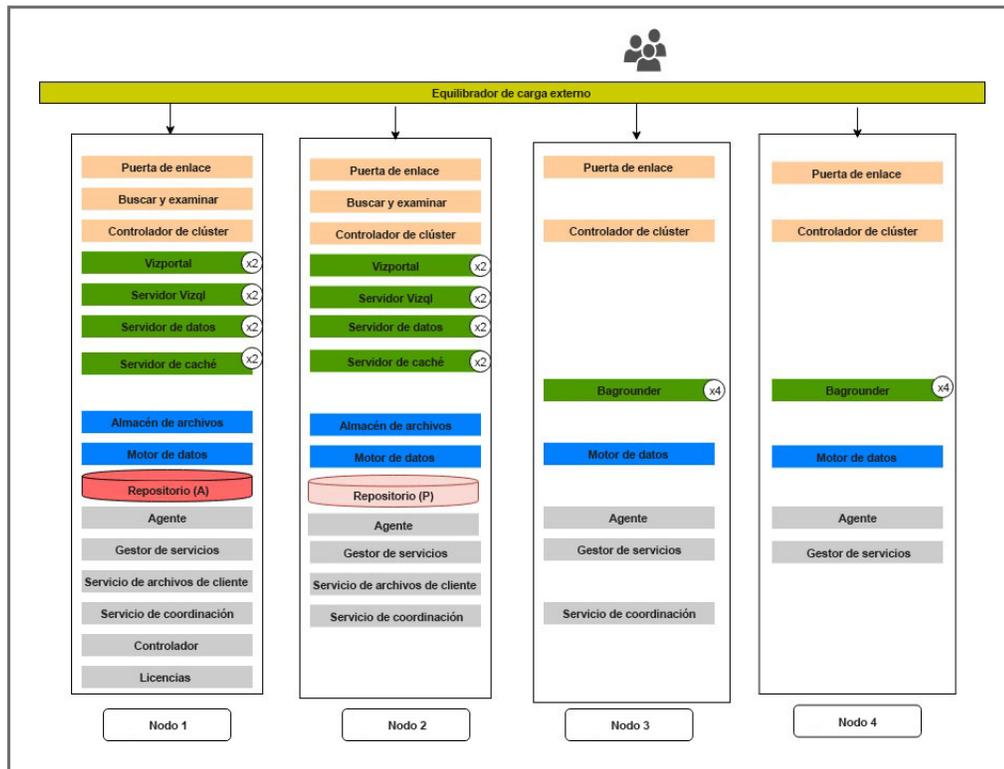


Figura 6 Ejemplo de una implementación básica de alta disponibilidad de 4 nodos.

El nodo inicial (Nodo 1) aloja las funciones administrativas y de licencia para todo el clúster. En esta configuración, si el Nodo 1 con el repositorio activo falla, el Nodo 3 con el repositorio pasivo se activa automáticamente. El nodo 1 y el nodo 2 en el clúster están configurados para potenciar la funcionalidad de visualización de datos de Tableau Server con redundancia en todo el clúster, y ambos nodos tienen múltiples instancias para VizPortal, VizQL Server, el servidor de datos y el servidor de caché. El nodo 3 y el nodo 4 están configurados como un nodo procesador en segundo plano dedicado. Esto asegurará que la carga de trabajo de la actualización de extracción no interfiera con la visualización de datos. El servicio de coordinación se implementa en los tres nodos. Si alguno de los nodos con el servicio de coordinación falla, todavía habría cuórum (dos de cada tres nodos siguen funcionando). Si el Nodo 4 y otro nodo del clúster fallan, el clúster continuará funcionando ya que el servicio de coordinación tendrá cuórum. Si dos nodos del nodo 1, el nodo 2 o el nodo 3 fallan, el clúster ya no tendrá quórum, ya que dos de los tres servicios de coordinación han fallado. En tal situación, necesitas un plan para recuperar tu sistema de este desastre.

Dejando a un lado otras consideraciones y simplemente centrándote en el supuesto de que deseas el menor tiempo de inactividad, debes considerar una arquitectura de implementación que sea de cinco o más nodos. Si bien recomendamos la arquitectura de 3 nodos en este documento como un gran punto de partida, las implementaciones grandes en toda la empresa con requisitos de alta disponibilidad cruciales deberían considerar el uso de cinco o más nodos como referencia.

Para las organizaciones que dependen en gran medida de las extracciones y las actualizaciones de la extracción, una arquitectura de implementación comúnmente utilizada incluye la implementación de procesos en segundo plano en los nodos de trabajo dedicados. Las cargas de trabajo de actualización de extracción a veces pueden interferir con las cargas de trabajo de visualización admitidas por los procesos de VizQL Server, por lo que al dividir los procesos en segundo plano en las máquinas de trabajo dedicadas, puedes evitar la contención de recursos entre estas dos cargas de trabajo diferentes. Para asegurarte de que tienes redundancia de proceso en varios nodos en el clúster, los nodos sólo de proceso en segundo plano a menudo se implementan en pares.

Conjunto de servicio de coordinación externa

El servicio de coordinación puede generar una gran cantidad de E/S cuando se comunica con otros componentes del servidor, por lo que si estás ejecutando Tableau Server en computadoras que cumplen o exceden los requisitos mínimos de hardware o simplemente deseas tener una instalación de Tableau Server más robusta, es posible que quieras instalar Tableau Server en una configuración que utilice nodos exclusivos del servicio de coordinación. Esto significa instalar el servicio de coordinación en nodos que no ejecutan otros procesos de servidor y eliminar el servicio de coordinación de cualquier nodo que esté ejecutando otros procesos de servidor. Consulta [configurar nodos exclusivos para el servicio de coordinación](#) para obtener más información sobre la configuración del conjunto de servicios de coordinación en hardware dedicado. La siguiente figura muestra una configuración de clúster con un conjunto de servicios de coordinación externos.

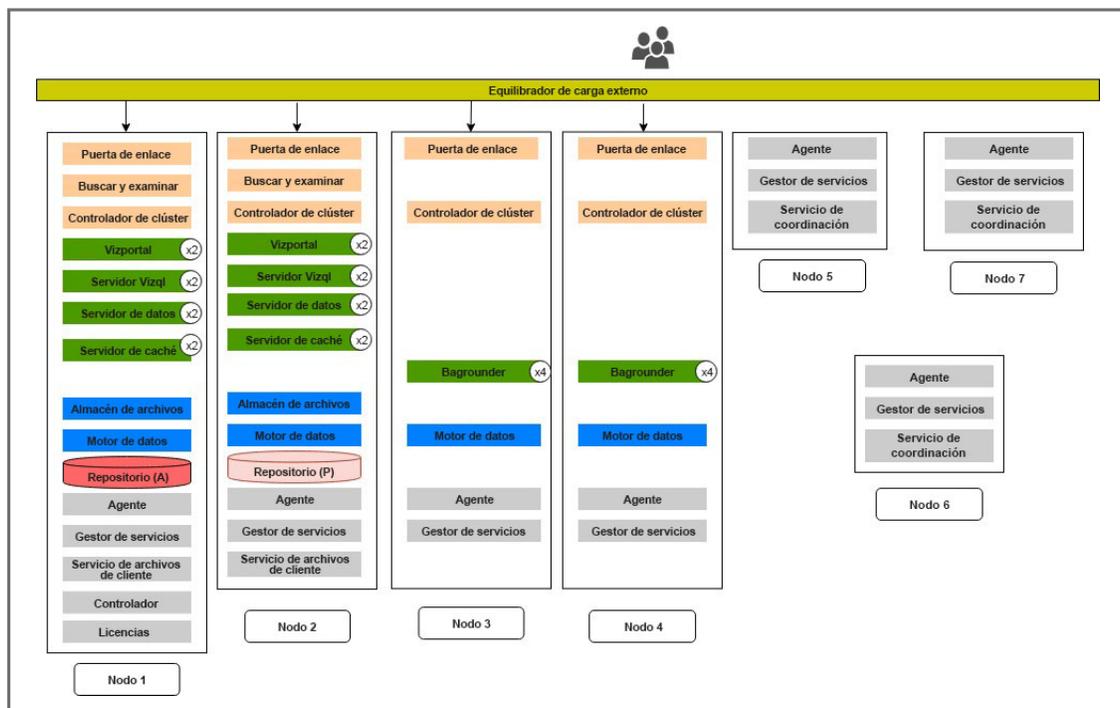


Figura 7 Ejemplo de una implementación de alta disponibilidad del servicio de coordinación externo.

Más allá de la alta disponibilidad

Lograr una alta disponibilidad requiere algo más que las características y funcionalidades integradas en Tableau Server. Los errores pueden ocurrir fuera del alcance del software de Tableau Server, como en el caso de errores de hardware o de red. Puedes lograr una alta disponibilidad mejorada al implementar Tableau Server en una infraestructura que sea más resistente a este tipo de errores, como el hardware virtualizado en nubes públicas o privadas.

La verdadera alta disponibilidad proviene de entender lo que necesitan los usuarios y seguir las prácticas y procesos recomendados para satisfacer esas necesidades. Aunque Tableau hace que la alta disponibilidad sea fácil de lograr con sus funcionalidades robustas, éstas no deben reemplazar las prácticas recomendadas, como el mantenimiento regular y las copias de seguridad programadas. Tableau es consciente de la importancia de las aplicaciones de inteligencia empresarial para ayudar a tu empresa a tomar decisiones. Tableau Server estará listo para las preguntas de análisis de tu organización.

Acerca de Tableau

Tableau es una plataforma de inteligencia de negocios visual completa, fácil de usar y preparada para la empresa, que ayuda a las personas a ver y comprender los datos con un análisis de autoservicio rápido y escalable. Ya sea en las instalaciones físicas o en la nube, en Windows o Linux, Tableau aprovecha las inversiones en tecnología existentes y se adapta a medida que el entorno de datos cambia y crece. Libera el poder de tus activos más valiosos: los datos y las personas.

Recursos adicionales

[Introducción a Tableau Services Manager](#)

[Instalaciones distribuidas y de alta disponibilidad de Tableau Server](#)

[Alta disponibilidad y recuperación de desastres en Tableau Server](#)

[Escalabilidad de Tableau Server](#)

