

Tableau Server 10.0

Alta disponibilidad: análisis imprescindibles a gran escala

Kitty Chou, administradora de productos
Mike Klaczynski, marketing de productos

Índice

| | |
|--|----|
| Análisis de autoservicio: una herramienta imprescindible..... | 3 |
| ¿Qué es la alta disponibilidad? | 3 |
| Escalabilidad de Tableau Server | 4 |
| Alta disponibilidad inmediata | 4 |
| ¿Qué es la alta disponibilidad de Tableau Server? | 5 |
| General | 5 |
| Puerta de enlace | 5 |
| Servidor de aplicación..... | 6 |
| Servicio de coordinación..... | 6 |
| Controlador de clúster | 7 |
| Repositorio | 7 |
| Componente de segundo plano | 9 |
| Servidor de datos | 9 |
| Servidor de caché..... | 9 |
| Motor de datos..... | 9 |
| Almacenamiento de archivos | 10 |
| Buscar y examinar | 11 |
| VizQL Server..... | 11 |
| Integración con herramientas de supervisión de terceros..... | 11 |
| Conmutación tras error de un nodo de servidor principal | 12 |
| Configuración de un servidor principal de seguridad..... | 13 |
| Supervisión del estado de los clústeres..... | 14 |
| Consideraciones sobre arquitectura..... | 15 |
| Selección de la mejor configuración | 15 |
| Implementación de alta disponibilidad con un mínimo de tres nodos..... | 16 |
| Implementaciones de más de tres nodos..... | 17 |
| Recuperación ante desastres con Tableau Server | 18 |
| La estrategia de recuperación ante desastres más simple | 18 |
| Soluciones de recuperación ante desastres de terceros..... | 18 |
| Capacidades de recuperación ante desastres de Tableau | 18 |
| Más allá de la alta disponibilidad | 19 |

Análisis de autoservicio: una herramienta imprescindible

Como nunca antes, el análisis de autoservicio y las decisiones basadas en datos son el estándar en organizaciones de todo el mundo. Para conseguir respuestas en tiempo real, los responsables de la toma de decisiones y los usuarios dependen del acceso inmediato a los datos y a las herramientas de autoservicio. Los ejecutivos comprenden la importancia de las decisiones basadas en datos para sus empresas y dependen de estos sistemas cada día. Esta dependencia de los datos exige un alto nivel de disponibilidad de los sistemas subyacentes. Es necesario que los equipos existentes y las herramientas empresariales tengan un mayor acceso a las funcionalidades de la plataforma y que estas se puedan configurar de manera sencilla.

Tableau Server 10.0 es el futuro del análisis de autoservicio para áreas fundamentales. Posibilita una exploración de datos rápida y de autoservicio, promueve la confianza en el contenido y los datos mediante una administración eficaz y es fácil de implementar, administrar y escalar en cualquier empresa. En este informe, exploraremos la forma en que Tableau Server 10.0 proporciona el análisis de autoservicio a gran escala con una alta disponibilidad.

¿Qué es la alta disponibilidad?

El objetivo de los sistemas de alta disponibilidad es minimizar sus períodos de inactividad. La disponibilidad se expresa comúnmente en “cantidad de números 9”, y se mide efectivamente como porcentaje de disponibilidad real frente a disponibilidad esperada. En la tabla siguiente, se muestra la correspondencia entre números 9 y los períodos de inactividad anuales.

| Cantidad de 9 | Porcentaje de disponibilidad | Total anual de tiempo de inactividad |
|---------------|------------------------------|--------------------------------------|
| 1 | 90 % | 36,5 días |
| 2 | 99 % | 3 días, 15 horas |
| 3 | 99,9 % | 8 horas, 45 minutos |
| 4 | 99,99 % | 52 minutos, 34 segundos |
| 5 | 99,999 % | 5 minutos, 15 segundos |

Figura 1: Métricas de disponibilidad típicas y períodos de inactividad correspondientes

A menudo, los administradores de sistema cuentan con acuerdos de nivel de servicios (SLA) con sus usuarios corporativos a fin de definir un umbral aceptable para los períodos de inactividad. A partir de dichos acuerdos se definen las arquitecturas de implementación para satisfacer los objetivos establecidos. En su mayoría, los administradores de sistema planean períodos de inactividad para realizar mantenimiento y hacer actualizaciones y revisiones. Además, es probable que haya fallos inesperados, habitualmente denominados períodos de inactividad no planeados. Desde luego, los administradores necesitan realizar mantenimiento planeado para actualizaciones de hardware o software; el objetivo es minimizar los períodos de inactividad no planeados.

Sabemos lo importante que es para los usuarios tener un fácil acceso a los datos para poderlos ver y comprender. También somos conscientes de que siempre habrá eventos que amenacen la disponibilidad de los sistemas de inteligencia de negocios, sin importar si están relacionados con hardware, software, redes o incluso personas. Como mínimo, los procesos de Tableau Server 10.0 se reiniciarán automáticamente para que su sistema siga funcionando en caso de fallo de un componente. Una implementación de nodos múltiples correctamente configurada emplea procesos redundantes para garantizar la alta disponibilidad del servidor. Sin embargo, a diferencia de la mayoría de los sistemas, Tableau facilita la instalación y configuración de su entorno de análisis para garantizar una alta disponibilidad.

Escalabilidad de Tableau Server

Tableau Server está diseñado para poder escalar en vertical y en horizontal. Es útil para las organizaciones grandes, puesto que proporciona estabilidad de implementación de nivel empresarial. Al mismo tiempo, es simple y fácil de usar, lo que es atractivo para equipos más pequeños. Según su entorno, Tableau Server puede ejecutarse en uno o más equipos (y ejecutar uno o más procesos de componente en el mismo nodo), a fin de satisfacer las exigencias de los usuarios y los requisitos de alta disponibilidad.

Internamente, Tableau administra varias implementaciones de Tableau Server en la escala de la nube. Entre ellas, se encuentran Tableau Public y Tableau Online. Tableau Public es una implementación personalizada de Tableau Server, que, todas las semanas, recibe millones de vistas públicas de todas partes del mundo. Como parte de nuestros procesos de ingeniería y lanzamiento, implementamos versiones beta de Tableau Server en Tableau Public para que, antes de ofrecérselas a nuestros clientes corporativos, podamos ajustar su estabilidad y calidad.

Alta disponibilidad inmediata

Es sencillo instalar y configurar la alta disponibilidad de Tableau Server. Una instalación predeterminada toma tan solo unos minutos. Puede instalar Tableau Server en el equipo principal, junto con todos los otros nodos de trabajador del clúster. A continuación, puede configurar y conformar un clúster de alta disponibilidad mediante la utilidad de configuración.

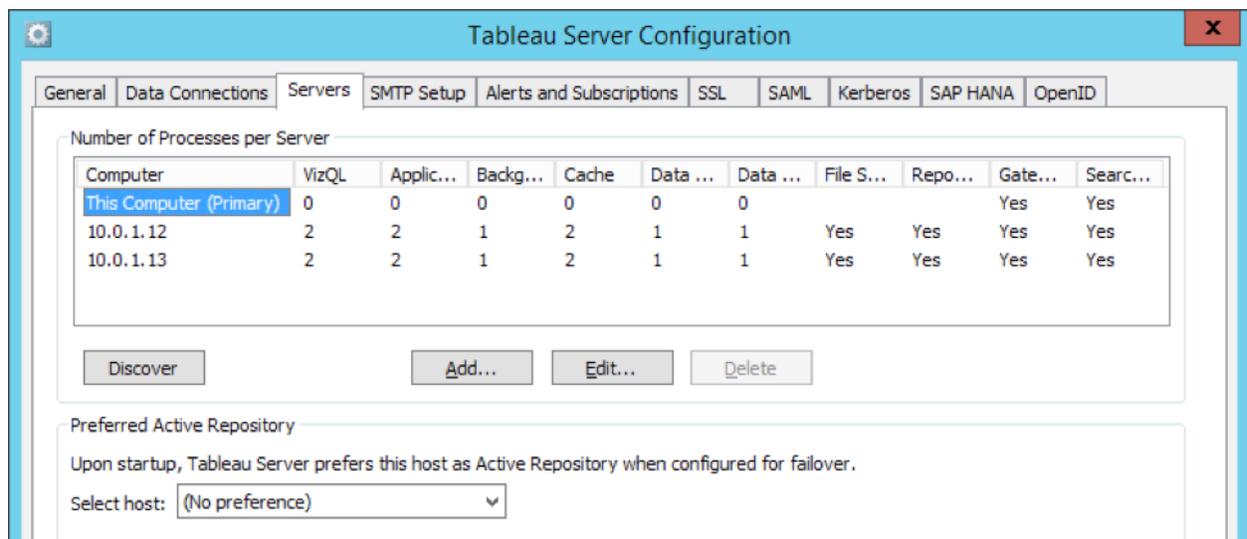


Figura 2: La utilidad de configuración de Tableau Server ofrece flexibilidad para contar y configurar los nodos y los procesos por nodo de un clúster.

¿Qué es la alta disponibilidad de Tableau Server?

Tableau Server tiene varios procesos dedicados que garantizan que los usuarios finales puedan acceder a todo el sistema. Para comprender esta sección es necesario conocer los componentes de Tableau Server y sus funciones. Si no está familiarizado con ellos, le sugerimos que, primero, lea la [Guía de administrador de Tableau Server](#).

Para saber cómo lograr la alta disponibilidad de Tableau Server, debe comprender cómo lograrla para cada uno de sus componentes. En otras palabras, se debe garantizar la alta disponibilidad de todos los componentes a fin de que el clúster completo de Tableau Server sea de alta disponibilidad. Esto se logra mediante la implementación de redundancias en puntos de error únicos. Analicemos cada componente por separado.

General

Tableau Server cuenta con procesos automáticos incorporados para reiniciar todos sus procesos de servidor. Gracias al reinicio automático de los procesos de servidor que fallan y su correspondiente notificación a los administradores, dicha automatización garantiza la alta disponibilidad. El hardware o la máquina virtual que hospede Tableau Server debe estar en condiciones para permitir la automatización.

A fin de evitar el fallo de todo el nodo, es importante configurar los procesos de servidor únicos, para generar redundancia en diferentes nodos del clúster. Es posible generar esa redundancia para todos los procesos, excepto para el servicio de licenciamiento, que solo puede ejecutarse en el nodo principal. Puede encontrar los detalles acerca de cómo manejar la limitación de este proceso en la sección *Conmutación tras error del servidor principal*.

Puerta de enlace

A partir de Tableau Server 8.1, el proceso de puerta de enlace puede ejecutarse en cualquier nodo del clúster de Tableau Server, no solo en el principal. Esto tiene repercusiones importantes para la alta disponibilidad.

En las versiones anteriores a Tableau Server 8.1, la ejecución del proceso de puerta de enlace solo se podía configurar en un solo nodo. Cuando ese proceso fallaba, Tableau Server no podía establecer comunicaciones fuera del clúster. Además, generaba alteraciones en muchas de las comunicaciones internas. En Tableau Server 10.0, cuando se configuran varios nodos para ejecutar el proceso de puerta de enlace, la comunicación con el servidor no sufre alteraciones, siempre que esté en ejecución por lo menos un proceso de puerta de enlace.

Mitigación de riesgos relacionados con fallos en la puerta de enlace

La clave para lograr una alta disponibilidad de la puerta de enlace es tener más de un nodo en el clúster de Tableau Server y configurarlos para que varios ejecuten el proceso de puerta de enlace. De hecho, le recomendamos que configure un proceso de puerta de enlace en cada nodo. Esto previene el riesgo de que el proceso de puerta de enlace sea un único punto de error que provoque que el servicio no esté disponible.

¿Qué sucede cuando falla un proceso de puerta de enlace? Como se mencionó anteriormente, si no hay ningún proceso de puerta de enlace en ejecución, no estará disponible el clúster de Tableau Server en su totalidad. Si hay otros procesos de puerta de enlace en ejecución, se les puede enviar las solicitudes a fin de que sean procesadas con normalidad. Sin embargo, cualquier solicitud recibida por la puerta de enlace fallida no podrá redireccionarse y continuará fallando, a pesar de haber otras puertas de enlace en funcionamiento. Los procesos de puerta de enlace fallidos se reinician automáticamente. Por lo tanto, siempre que el equipo esté en funcionamiento, el proceso fallido se ejecutará nuevamente y volverá a procesar las solicitudes.

A fin de que su instalación sea incluso más eficaz ante los fallos de puerta de enlace, le recomendamos ubicar el clúster de Tableau Server bajo el control de un equilibrador de carga externo. De esta manera, garantizará que las solicitudes solo se dirijan a procesos de puerta de enlace en funcionamiento. En la [Guía de administrador de Tableau Server](#), encontrará información adicional sobre cómo configurar esta opción.

Servidor de aplicación

Es sencillo lograr una alta disponibilidad del servidor de aplicación. Simplemente debe configurar instancias del servidor de aplicación en todos los nodos del clúster de Tableau Server.

¿Qué sucede cuando falla un proceso de servidor de aplicación? Fallarán las solicitudes que maneje esa instancia. Sin embargo, las solicitudes siguientes se dirigirán a otros procesos de servidor de aplicación en ejecución. Si se sigue ejecutando el nodo que contiene el proceso fallido de servidor de aplicación, el proceso fallido debería reiniciarse en breve.

Servicio de coordinación

El servicio de coordinación es parte de la instalación de base de Tableau Server. La cantidad de nodos instalados con el servicio de coordinación depende de la cantidad de nodos del clúster, según se detalla en la tabla a continuación:

| Cantidad de nodos en el clúster | Cantidad de nodos con servicios de coordinación |
|---------------------------------|---|
| 1 - 2 | 1 |
| 3 - 4 | 3 |
| 5 o más | 5 |

El proceso de servicio de coordinación se instalará de manera secuencial en los primeros n nodos del clúster (incluido el nodo principal); n es la cantidad de nodos con servicio de coordinación (según la tabla a continuación).

¿Qué es el cuórum?

Si la cantidad de procesos de servicio de coordinación no llega al cuórum, Tableau Server experimentará una interrupción total. El cuórum se calcula sobre la cantidad total de procesos de servicio de coordinación configurados. Cuórum es lo mismo que mayoría absoluta. Por ejemplo, un clúster con tres o cuatro equipos tiene la capacidad de tolerar la pérdida de un nodo (una instancia del servicio de coordinación) como máximo. Un clúster con cinco o seis equipos puede tolerar la pérdida de hasta dos procesos de servicio de coordinación.

En particular, un clúster con solo dos nodos no puede tolerar la pérdida de ni siquiera un único proceso de servicio de coordinación. Esta es la razón por la que, para lograr la alta disponibilidad total (incluida la conmutación tras error automática), se requiere un mínimo de tres nodos.

¿Qué sucede cuando falla un proceso de servicio de coordinación? Nada, siempre y cuando la cantidad de procesos de servicio de coordinación siga alcanzando el cuórum. Si la cantidad de procesos de servicio de coordinación en funcionamiento es menor que el cuórum, todo el clúster de Tableau Server dejará de estar disponible para proteger la integridad referencial de la base de datos Postgres subyacente.

Los procesos fallidos de servicio de coordinación se reinician automáticamente siempre que el equipo funcione de manera correcta.

Controlador de clúster

El controlador de clúster también es parte de la instalación de base de Tableau Server 10.0. Durante la instalación, se configura un proceso de controlador de clúster en cada nodo del clúster, sin necesidad de configuración explícita.

¿Qué sucede cuando falla un proceso de controlador de clúster? Todos los otros componentes de Tableau Server del mismo nodo dejarán de estar disponibles y aparecerán como “No disponible” en la página de estado de Tableau Server. Tampoco estará disponible ningún proceso de repositorio que se esté ejecutando en ese nodo. (Para saber más acerca de este efecto, consulte la sección *Repositorio* a continuación). Para evitar fallos del proceso de controlador de clúster, asegúrese de que haya redundancia para todos los componentes de servidor únicos y que estos se ejecuten, al menos, en dos nodos diferentes del clúster.

Los procesos fallidos de controlador de clúster se reinician automáticamente siempre que el equipo funcione de manera correcta. Cuando se reinicia el controlador de clúster, también se reinician todos los procesos de repositorio configurados en ese nodo.

Repositorio

El repositorio es una base de datos necesaria para que Tableau Server funcione. Si no se cuenta con un repositorio que funcione correctamente, todo el clúster de Tableau Server no estará disponible. Solo puede haber en uso un único repositorio a la vez. Se lo denomina repositorio “activo”. Todas las operaciones que requieran el repositorio se ejecutan en este repositorio activo.

Para mejorar la disponibilidad, puede configurar Tableau Server con un repositorio adicional, denominado “pasivo”, en un nodo diferente del clúster. Los contenidos del repositorio activo se transfieren constantemente al repositorio pasivo. Si falla el repositorio activo, un clúster con una configuración de alta disponibilidad cambiará automáticamente el estado del repositorio pasivo a activo a fin de garantizar la disponibilidad del servidor. Para los clientes que deseen una alta disponibilidad, es recomendable configurar un repositorio pasivo. En todo un clúster, solo puede haber dos repositorios (uno activo y otro pasivo) como máximo. No pueden estar ubicados en el mismo nodo.

El controlador de clúster administra el inicio y el apagado del repositorio, además de cualquier conmutación tras error que requiera un cambio de estado de pasivo a activo. Esta es la razón por la que un fallo en el proceso de controlador de clúster puede causar un problema con el repositorio. Si falla el proceso de controlador de clúster que inició el repositorio, este último también fallará.

¿Qué sucede cuando falla un proceso de repositorio? Depende de varios factores. En la lista siguiente se resumen los diversos casos:

- Si falla el repositorio pasivo, los usuarios no deberían verse afectados. Todo continuará normalmente, puesto que el repositorio activo sigue en funcionamiento. En segundo plano, se reiniciará el repositorio pasivo y se reanudará la replicación de datos. Sin embargo, es posible que el repositorio pasivo tarde un tiempo en volver a sincronizarse con el activo.

- Si falla el repositorio activo y si no hay un repositorio pasivo completamente sincronizado, Tableau Server no estará disponible hasta que se reinicie el repositorio activo. El sistema intentará hacerlo automáticamente, pero, según el motivo del fallo, es posible que esto no suceda.

Nota: Esto implica que, si actualmente solo se puede utilizar el repositorio activo, la disponibilidad de Tableau Server no se puede considerar alta. El repositorio activo es un punto de error único para todo el sistema si no hay un repositorio pasivo sincronizado.

- Si falla el repositorio activo, hay disponible un repositorio pasivo completamente sincronizado y el clúster está configurado para alta disponibilidad, se ejecutará automáticamente la conmutación tras error para utilizar el repositorio pasivo. Después de la conmutación tras error, el repositorio que antes era pasivo pasará a ser el nuevo repositorio activo. El sistema reiniciará el repositorio fallido, anteriormente activo, y lo utilizará como el nuevo repositorio pasivo. A continuación, se iniciará la sincronización. También reiniciará otros procesos pertinentes de manera automática, para que reconozcan el nuevo repositorio activo y se puedan volver a conectar. Durante este breve período de reinicios, los usuarios experimentarán una interrupción del servicio. Sin embargo, los administradores no deben intervenir para asegurar una continuidad en la disponibilidad, puesto que esto es una secuencia automática. Si desea establecer el repositorio pasivo como activo de manera manual, lo puede hacer con el comando `tabadmin failoverrepository`.

Componente de segundo plano

Para proporcionar alta disponibilidad al servicio de un componente de segundo plano, debe configurar el clúster de manera que se ejecute más de un proceso de segundo plano en varios nodos. Cuando deba

decidir dónde y cuántos componentes de segundo plano se ejecutarán, tenga en cuenta la capacidad disponible de cada equipo sobre la base de los otros procesos de servidor.

¿Qué sucede si deja de ejecutarse un proceso de componente de segundo plano? Fallarán las tareas en las que esté trabajando el componente de segundo plano y no se intentará completarlas nuevamente. Sin embargo, la mayoría de las tareas de segundo plano se ejecutan de manera periódica según un programa. Un proceso de componente de segundo plano en funcionamiento volverá a ejecutar normalmente las mismas tareas en el próximo turno programado.

Los procesos fallidos de componente de segundo plano se reinician de manera automática siempre que el equipo funcione correctamente, pero no se volverán a ejecutar las tareas que fallaron.

Servidor de datos

Con el propósito de asegurar una alta disponibilidad del servidor de datos, configure uno o más procesos de servidor de datos para que se ejecuten en varios nodos del clúster.

¿Qué sucede cuando falla un proceso de servidor de datos? Fallarán las consultas que se ejecuten mediante el proxy a través del proceso de servidor de datos y la vista no se representará correctamente. Las solicitudes subsiguientes, incluido un reintento de la operación fallida, deberían ejecutarse con éxito, siempre que haya un servidor de datos en funcionamiento y que pueda aceptar solicitudes redireccionadas.

Para funcionar, Tableau Server no depende del servidor de datos. Sin embargo, sin un servidor de datos en funcionamiento, el clúster pierde su capacidad de enviar los libros de trabajo a través del proxy hasta las fuentes de datos externas. Las vistas que no requieran un servidor de datos para una de sus fuentes de datos deberían funcionar correctamente.

Servidor de caché

El servidor de caché proporciona una memoria caché de consultas externa y compartida. Se trata de una caché de pares clave/valor que almacena información de consultas previas y acelera las solicitudes futuras. Esto implica que, si un proceso de servidor de caché deja de estar disponible (o, incluso, si ninguno está disponible), las consecuencias son relativamente leves. Tableau Server seguirá funcionando, aunque es posible que las acciones demoren más, puesto que los resultados almacenados previamente en la memoria caché no están disponibles. A medida que se vuelven a ejecutar las consultas, se envían nuevamente al servidor de caché reiniciado, lo que, con el tiempo, acelerará el proceso para los usuarios finales. De hecho, el servidor de caché no tiene un impacto en la disponibilidad. Sin embargo, tiene un impacto efectivo para el usuario final en diversas situaciones. Al igual que todos los demás procesos, el servidor de caché se reiniciará automáticamente, siempre que el equipo funcione correctamente.

Motor de datos

Cuando se utiliza el análisis en memoria, el componente de motor de datos carga las extracciones de datos y les envía consultas. Con el propósito de asegurar una alta disponibilidad del motor de datos, configure uno o más procesos de motor de datos para que se ejecuten en varios nodos del clúster. Todos los motores de datos se ejecutan en modo activo/activo y realizan exactamente las mismas funciones. Es importante tener

en cuenta que todos los nodos configurados para que ejecuten un proceso de motor de datos también lo estarán para ejecutar el proceso de almacenamiento de archivos. El proceso de almacenamiento de archivos administra el almacenamiento y la replicación de los archivos de extracción y se describe a continuación.

¿Qué sucede si deja de ejecutarse un proceso de motor de datos? Fallarán las consultas actualmente en ejecución en ese proceso de motor de datos, lo que dará como resultado un error en la representación de vista o en la actualización de la extracción. Si la misma operación se vuelve a ejecutar, se reasignará automáticamente a un motor de datos diferente que esté en funcionamiento.

Los procesos fallidos de motor de datos se reinician automáticamente, siempre que el equipo funcione de manera correcta.

Almacén de archivos

Como se mencionó anteriormente, un proceso de almacén de archivos se iniciará de manera automática en todos los nodos que ejecuten uno o más procesos de motor de datos. El proceso de almacén de archivos administra el almacenamiento y la replicación de archivos de extracción entre nodos.

¿Cómo funciona el almacén de archivos?

Un archivo de extracción se crea en el sistema cuando un usuario publica uno por primera vez en Tableau Server o cuando se actualiza. Inmediatamente después de alguno de estos eventos, la extracción solo se puede encontrar en un único almacén de archivos de un único nodo. Puesto que el archivo de extracción no se puede considerar de alta disponibilidad ya que no se configuró su redundancia, constituye un único punto de error. Los procesos de almacén de archivos se comunican entre sí para replicar rápidamente las extracciones locales en todos los demás nodos de almacén de archivos del clúster. Siempre que los recursos de la red del clúster lo permitan, este proceso está diseñado para copiar los archivos rápidamente. Sin embargo, el tiempo que tarde dependerá del tamaño de la extracción. En cuanto una copia está disponible en varios nodos de los clústeres, se puede considerar el archivo de extracción como de alta disponibilidad.

¿Qué sucede si falla un proceso de almacén de archivos? Tiene dos consecuencias:

- No se podrá copiar archivos de extracción desde el nodo afectado ni en él.
- No se podrá eliminar los archivos de extracción que ya no sean necesarios del nodo afectado. Este proceso de eliminación se podría considerar como una especie de “purga de extracciones”.

La suspensión de la purga de extracciones no tiene efectos inmediatos. Simplemente significa que se acumularán archivos de extracción no deseados que ocuparán espacio en disco en ese nodo. Con el tiempo, esto podría generar un problema, aunque los nodos con el tamaño adecuado deberían tener un búfer amplio de espacio en disco.

Cuando no se produce la replicación, los archivos de extracción recientemente agregados al nodo fallido de almacén de archivos no estarán disponibles en ningún otro nodo del mismo tipo en el clúster y viceversa. En cuanto se reinicie el proceso de almacén de archivos, el sistema se autocorregirá mediante la sincronización de todos los almacenes de archivos en todos los nodos.

Los procesos fallidos de almacén de archivos se reinician automáticamente siempre que el equipo funcione de manera correcta. El proceso de almacén de archivos debería reanudar sus funciones rápidamente, como la sincronización de todos los archivos, independientemente de si fueron agregados durante el período de error o después.

Buscar y examinar

Es sencillo lograr una alta disponibilidad para el proceso de buscar y examinar. El sistema se debe configurar de modo que el proceso de buscar y examinar se ejecute en varios equipos.

¿Qué sucede si falla un proceso de buscar y examinar? En su mayor parte, Tableau Server no se podrá utilizar y, a pesar de que los usuarios podrán iniciar sesión en el sistema, parecerá que los libros de trabajo estarán vacíos de contenido. En realidad, no es que estén vacíos. El sistema simplemente no puede recuperar el contenido en los resultados de la búsqueda. Se mostrarán una vez que se haya reiniciado el proceso de buscar y examinar. Si, en el momento del fallo, hay más de un proceso de buscar y examinar configurado y se están ejecutando en varios nodos, también fallarán las solicitudes enviadas a un proceso de este tipo con error. Sin embargo, las solicitudes subsiguientes se redirigirán a un proceso de buscar y examinar en funcionamiento. Todos los procesos de buscar y examinar llevan un registro de índices entre todos los nodos del clúster. Por lo tanto, si un solo proceso de este tipo sigue funcionando (a pesar de que el resto no lo haga), aún se devolverán resultados para todos los nodos.

VizQL Server

Para lograr una alta disponibilidad del proceso de VizQL Server, simplemente configúrelo de modo que por lo menos una instancia se ejecute en varios equipos.

¿Qué sucede cuando falla un proceso de VizQL Server? Si solo hay un proceso de VizQL Server y falla, Tableau Server no podrá generar ninguna vista. La alta disponibilidad requiere la configuración de procesos de VizQL redundantes. Una configuración típica consta de entre dos y cuatro procesos de VizQL Server en cada nodo. Esto satisface simultáneamente las necesidades de alta disponibilidad y escalabilidad. Si hay varios procesos de VizQL Server en ejecución, el fallo de un solo proceso producirá solicitudes con error y la pérdida de datos de sesión en el momento del fallo. Cualquier solicitud futura será redirigida a otro proceso de este tipo en ejecución en el clúster de Tableau Server.

Hasta ahora, hemos visto cómo se comportan los procesos de servidor cuando fallan, además de la forma en que puede mitigar los riesgos de fallo a fin de garantizar la alta disponibilidad de todo el clúster de Tableau Server. Junto con el planeamiento de medidas para cada situación de fallo, también debería supervisar el clúster de manera proactiva para evitar errores en los procesos ocurridos con anterioridad.

Integración con herramientas de supervisión de terceros

Además de la supervisión del sistema mediante los mecanismos incorporados de Tableau Server, también puede recibir de manera remota un archivo XML de lectura mecánica con el estado de todos los procesos. Deberá habilitar el acceso remoto al servidor y utilizar la URL adecuada para su instancia: `http://<my_tableau_server>/admin/systeminfo.xml`

Recibirá un archivo `status.xml` que se puede analizar para identificar el estado del servidor y se puede integrar a otros sistemas o herramientas de supervisión:

Muestra del código XML de lectura mecánica del archivo `status.xml` proporcionado por acceso remoto a la URL de estado de servidor

```
<systeminfo xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <machines>
    <machine name="my_tableau_server">
      <repository worker="my_tableau_server:8060" status="Active"
preferred="false"/>
      <dataengine worker="my_tableau_server:27042" status="Active"/>
      <applicationserver worker="my_tableau_server:8600" status="Active"/>
      <apiserver worker="my_tableau_server:8000" status="Active"/>
      <vizqlserver worker="my_tableau_server:9100" status="Active"/>
      <dataserver worker="my_tableau_server:9700" status="Active"/>
      <backgrounder worker="my_tableau_server:8250" status="Active"/>
      <gateway worker="my_tableau_server:80" status="Active"/>
      <searchandbrowse worker="my_tableau_server:11000" status="Active"/>
      <cacheserver worker="my_tableau_server:6379" status="Active"/>
      <filestore worker="my_tableau_server:9345" status="Active"
pendingTransfers="0" failedTransfers="0" syncTimestamp="2015-02-
27T20:30:48.564Z"/>
      <clustercontroller worker="my_tableau_server:12012"
status="Active"/>
      <coordination worker="my_tableau_server:12000" status="Active"/>
    </machine>
  </machines>
  <service status="Active"/>
</systeminfo>
```

Conmutación tras error del nodo de servidor principal

El nodo principal de Tableau Server se identifica con el servidor en el que se haya instalado por primera vez la herramienta. El servidor principal es único, ya que incluye un único proceso de administración de licencias y otras funciones administrativas, y puede ser una instalación de servidor completamente autónoma. En implementaciones más grandes, el servidor principal se puede utilizar solo con las funciones administrativas básicas, lo que permite omitir procesos redundantes en los nodos de trabajador. Aunque Tableau Server proporciona una gran flexibilidad para las implementaciones, es importante recordar que sus procesos consultan el servidor principal cada 72 horas para realizar verificaciones de licencias. Si el servidor principal no está disponible en el momento de la verificación de licencias, se producirá un error y la implementación de Tableau Server pasará a estar “sin licencia” y “deshabilitada”.

Sin embargo, en el mundo real, su servidor principal puede fallar por numerosos motivos ajenos al dominio del software de Tableau Server; por ejemplo, un fallo del hardware, el SO o la red. En dichos casos, si implementó el clúster bajo el control de un equilibrador de carga externo y tiene puertas de enlace en todos los nodos, el resto del clúster seguirá estando disponible para responder solicitudes. Sin embargo, si el nodo

principal sigue estando deshabilitado en el momento en que el período de verificación de licencias vence, el clúster completo pasará a estar sin licencia. Como precaución para estos errores, asegúrese de contar con un respaldo del nodo principal en espera semiactiva.

Configuración de un servidor principal de seguridad

Si el servidor principal falla, se producirá un período de inactividad de las tareas administrativas hasta que se resuelva el problema que afecta el funcionamiento del servidor. Por esta razón, debe tener preparado un equipo de respaldo especialmente dedicado para funcionar como “copia de seguridad” del servidor principal en el caso de que este deje de operar correctamente. Una vez configurado y listo, el servidor principal de seguridad no se debe encender ni conectar al clúster. Esto garantiza que las funciones de licenciamiento y administrativas continúen funcionando con el nodo del servidor principal. Aprenda a configurar y preparar un servidor principal de seguridad con las siguientes instrucciones de nuestra [Guía de administrador de Tableau Server](#).

En caso de fallo del servidor principal, en nuestra [Guía de administrador de Tableau Server](#) también encontrará los pasos que debe seguir para cambiar al servidor principal de seguridad. Estos pasos sirven para enviar alertas al resto del clúster acerca del cambio y redirigir automáticamente las solicitudes administrativas y de licencias al ahora nuevo servidor principal. Estos pasos también se pueden automatizar.

Supervisión del estado de los clústeres

Los administradores de sistema tienen la capacidad de supervisar el estado de los procesos de Tableau Server 10.0 en la página de estado de Tableau Server. Proporciona a los administradores una visión general del funcionamiento del clúster, ya que muestra los procesos de servidor de todos los nodos, además del estado de los repositorios activo y pasivo.

Server Status

Process Status

The real-time status of processes running in Tableau Server.

| Process | Primary 10.0.1.11 | Worker 1 10.0.1.12 | Worker 2 10.0.1.13 |
|--------------------|----------------------|-----------------------|-----------------------|
| Cluster Controller | ✓ | ✓ | ✓ |
| Gateway | ✓ | ✓ | ✓ |
| Application Server | | ✓ ✓ | ✓ ✓ |
| VizQL Server | | ✓ ✓ | ✓ ✓ |
| Cache Server | | ✓ ✓ | ✓ ✓ |
| Search & Browse | ✓ | ✓ | ✓ |
| Backgrounder | | ✓ | ✓ |
| Data Server | | ✓ | ✓ |
| Data Engine | | ✓ | ✓ |
| File Store | | 🔄 Synchronizing | 🔄 Synchronizing |
| Repository | | ✓ | ✓ |

  Active  Busy  Passive  Unlicensed  Down Status unavailable

Figura 3. En la página de estado de servidor, los administradores pueden ver el estado de todos los procesos de cada nodo.

Además, mediante la utilidad de configuración, se pueden ajustar las opciones de Tableau Server para enviar una alerta de correo electrónico a los administradores del servidor en el momento en que se produzca un error del sistema. Esta utilidad también se puede utilizar a fin de proporcionar una advertencia temprana de problemas de capacidad de espacio en disco.

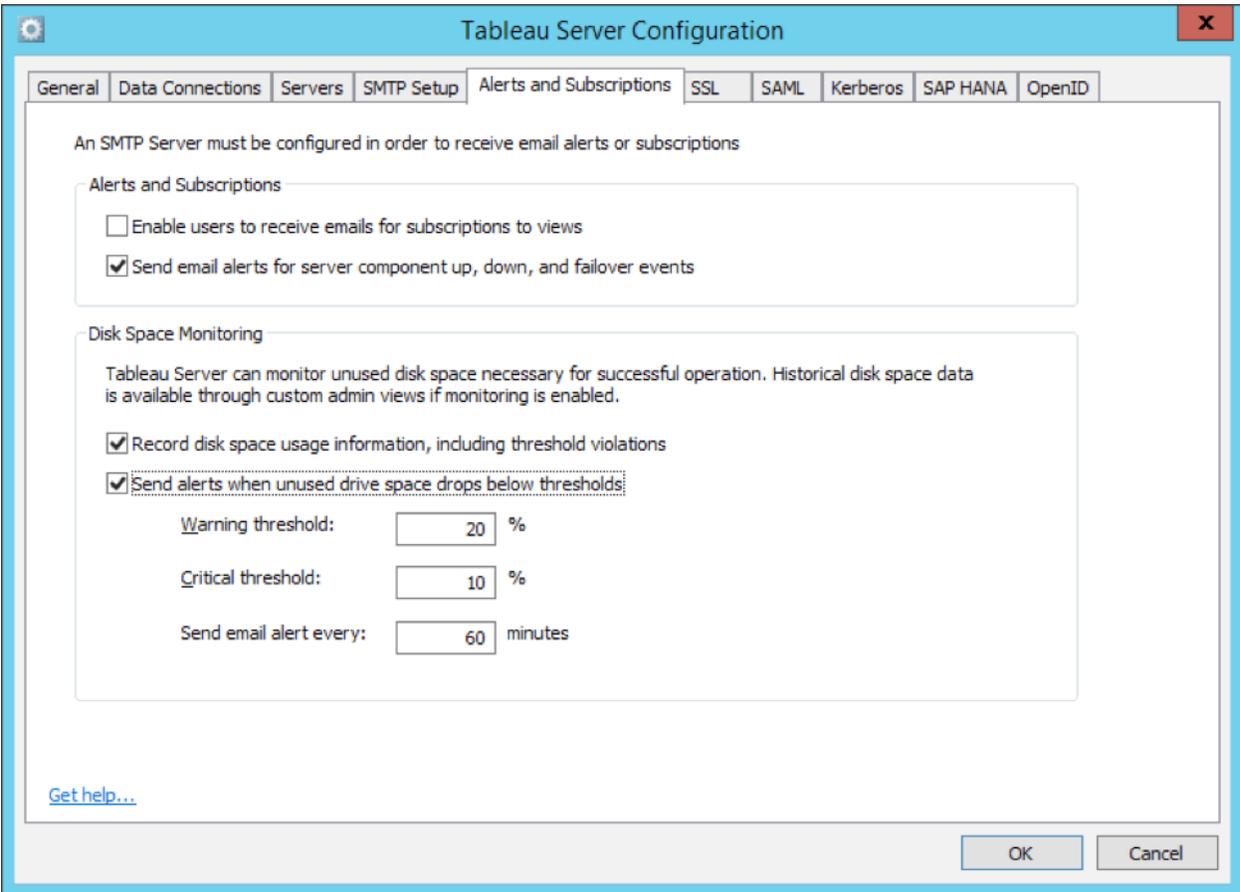


Figura 4. La utilidad de configuración de Tableau Server permite establecer alertas por errores del sistema.

Consideraciones sobre arquitectura

Para una implementación exitosa de la alta disponibilidad, debe comprender los objetivos de disponibilidad y los niveles de servicio que se esperan en su organización, además de planear la redundancia de manera correspondiente.

Selección de la mejor configuración

Las necesidades de su negocio y los recursos disponibles en su organización determinarán la configuración óptima de clústeres. Aunque hay diversas configuraciones posibles para lograr con éxito una alta disponibilidad, es importante conocer las características de los clústeres de diferentes tamaños a fin de poder elegir la mejor opción para su entorno. Además, puede elegir entre configuraciones simétricas (todos los nodos de trabajador tienen los mismos componentes) o asimétricas (todos los nodos de trabajador tienen cantidades y tipos variables de componentes). A menudo, la configuración simétrica para los nodos de trabajador simplifica la clonación y la incorporación de nodos adicionales al clúster. Sin embargo, mientras planea su configuración, tenga en cuenta las limitaciones de tener un solo repositorio activo y un solo repositorio pasivo.

Implementación de alta disponibilidad con un mínimo de tres nodos

Como se mencionó anteriormente, debe tener por lo menos tres nodos en el clúster para lograr una configuración de alta disponibilidad. Un clúster de tres nodos es también un punto de inicio excelente para implementaciones grandes. Si uno falla, todavía llega al cuórum con los dos nodos restantes. A continuación, en la figura 5, se muestran las responsabilidades del nodo principal en relación con las funciones administrativas y de licenciamiento. También se incluyen los dos nodos de trabajador adicionales que ejecutan las funcionalidades de datos y visualización de Tableau Server. La combinación exacta de estos componentes puede variar según sus necesidades de escalabilidad y alta disponibilidad. Por ejemplo, aunque en el gráfico se muestra solo un proceso de VizQL Server por nodo de trabajador, en realidad, usted podría hacer que se ejecuten dos o más de estos procesos en un nodo de trabajador para cubrir las necesidades de alta disponibilidad y las de escalabilidad. Además, si tiene más capacidad de procesamiento en su nodo principal, también puede incluir procesos de servidor adicionales, como VizQL Server o la combinación motor de datos/almacén de archivos, para una mayor escalabilidad.

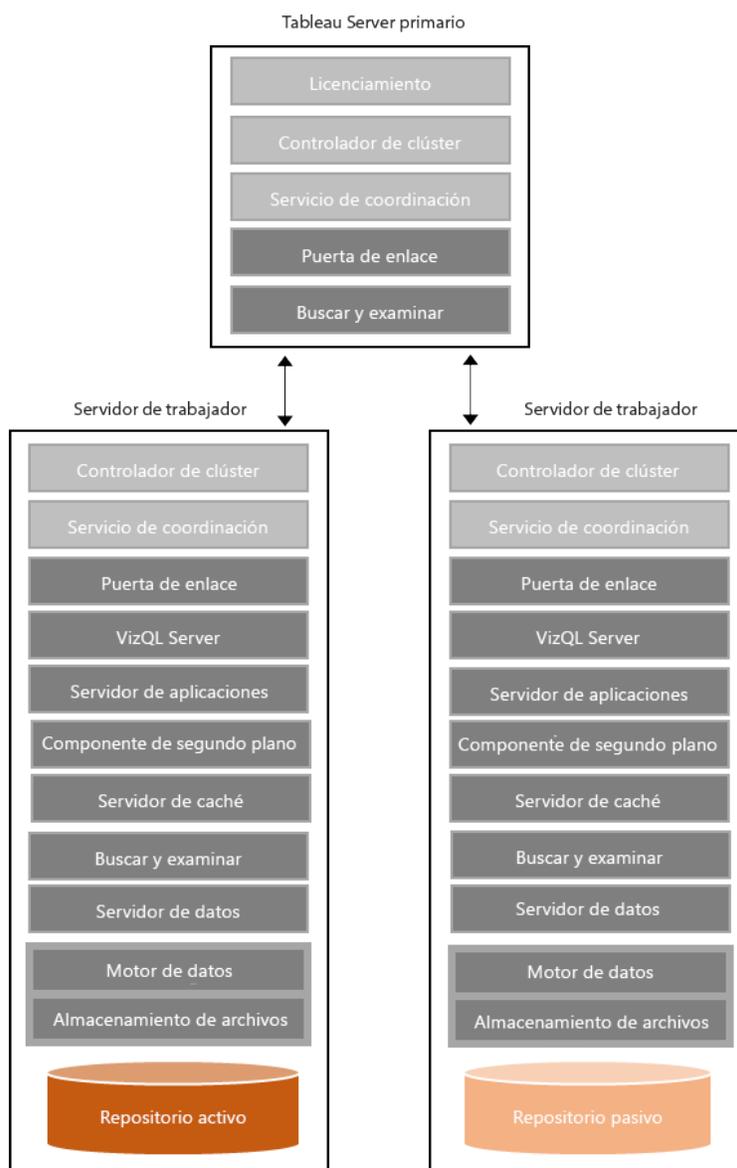


Figura 5. Ejemplo de implementación de alta disponibilidad con un mínimo de tres nodos

Al tener motores de datos instalados en diferentes equipos, cuenta con redundancia adicional en caso de que falle el hardware y se necesite configurar una rápida recuperación de copias de seguridad. Si necesita redundancia extra, puede implementar un motor de datos en cada nodo de trabajador.

Los procesos de su servidor principal no se verán afectados si algún nodo de trabajador falla. Sin embargo, los nodos de trabajador dependen del servidor principal. En esta configuración, si falla un nodo de trabajador con el repositorio activo, el nodo de trabajador del repositorio pasivo pasa a ser automáticamente el activo. Sin embargo, como se mencionó anteriormente, lo mejor es contar con un servidor principal de seguridad e instalarlo para que la recuperación del error en el servidor principal esté garantizada. Además, si su organización tiene acceso a un equilibrador de carga externo, debe implementar el clúster bajo el control de dicho dispositivo a fin de mejorar la disponibilidad para el usuario final. Si se está iniciando en el terreno de la alta disponibilidad, esta es una buena configuración para comenzar.

Implementaciones de más de tres nodos

Cuando utilice tres o más nodos, la topología recomendada es la implementación de una cantidad impar de nodos en un clúster para lograr una alta disponibilidad mejorada. Los clústeres con nodos pares tienen la misma capacidad de quórum que los clústeres con un nodo menos. Por ejemplo, en el caso de que se interrumpa el funcionamiento de dos nodos, las configuraciones con tres o cuatro nodos se verán afectadas de la misma manera, puesto que no alcanzarán el quórum y, por lo tanto, se producirá un apagado total.

La incorporación de un cuarto nodo a un clúster de tres para hacerlo par distribuirá la carga de trabajo entre más nodos, lo que ayuda a la escalabilidad. Además, reducirá el riesgo de un fallo total del clúster gracias a una mayor cantidad de procesos redundantes adicionales.

Dejando de lado otros factores y centrándonos en el supuesto de que prefiere la menor cantidad de tiempo de inactividad, debería considerar una arquitectura de implementación con cinco o más nodos. En este informe se recomienda la arquitectura de tres nodos como un excelente punto de partida. Sin embargo, para las implementaciones grandes, que abarcan toda la organización, y que tienen requisitos de alta disponibilidad para procesos fundamentales, se debería considerar una configuración de por lo menos cinco nodos como base.

En las organizaciones que dependen de extracciones y de sus actualizaciones se suele utilizar una arquitectura de implementación con procesos de componente de segundo plano en nodos de trabajador dedicados. En ocasiones, las cargas de trabajo resultantes de la actualización de extracciones interfieren con las cargas de trabajo de las visualizaciones producidas por los procesos de VizQL Server. Por este motivo, si distribuye los procesos de componente de segundo plano en equipos de trabajador dedicados, puede evitar que esas dos cargas de trabajo diferentes retengan recursos. Con el objetivo de garantizar la redundancia de procesos entre diversos nodos del clúster, es habitual que los nodos de trabajador que solo ejecutan procesos de componente de segundo plano se implementen en pares.

Como ya mencionamos, la alta disponibilidad reduce los tiempos de inactividad causados por fallos críticos. A pesar de satisfacer los requisitos de nivel de servicio e implementar una topología que permita una alta disponibilidad, es posible que aún se deba enfrentar a situaciones de error que exijan una recuperación rápida. Ante una situación de desastre o de error del hardware, debe planear la recuperación del sistema según sus objetivos de recuperación ante desastres.

Recuperación ante desastres con Tableau Server

Cuando planea la recuperación ante desastres en su entorno de Tableau, hay dos factores principales que debe tener en cuenta en relación con su negocio: (1) la tolerancia ante los períodos de inactividad que forman parte del proceso de recuperación y (2) la tolerancia a la pérdida de datos. Esto se define en términos del objetivo de tiempo de recuperación (RTO), una medida del tiempo que puede transcurrir antes de la recuperación total, y del objetivo de punto de recuperación (RPO), una medida de la tolerancia a la pérdida de datos.

Al determinar sus umbrales de RTO y RPO, es importante que comprenda los compromisos, que son principalmente financieros. Independientemente de las cuestiones específicas, un objetivo de tiempo de recuperación muy corto es costoso, debido a la infraestructura y el personal necesarios para volver a poner el sistema en funcionamiento con rapidez. Si opta por invertir menos en infraestructura y capacidad de respuesta del personal, debería ser capaz de tolerar un RTO más extenso (mayor tiempo de inactividad). Los compromisos no son específicos de Tableau Server y son aplicables a cualquier implementación de software de servidor.

Su tolerancia a la pérdida de datos se verá determinada por la aplicación individual y el índice de cambio de datos de la aplicación, entre otros aspectos. Por ejemplo, un RPO de 24 horas implica que, como máximo, solo perderá los datos de las últimas 24 horas.

La estrategia de recuperación ante desastres más simple

En muchos casos, la estrategia de recuperación ante desastres más simple y eficaz es realizar copias de seguridad nocturnas y contar con un equipo de servidor en espera, semiactiva o activa, instalado y listo en el centro de datos de recuperación. A diferencia de otras herramientas de BI tradicionales, Tableau Server es rápido de instalar, configurar y poner en marcha. Además, en la mayoría de los casos, puede alcanzar los objetivos de su negocio con la instalación de un servidor nuevo en un equipo en espera y hacer una restauración desde una copia de seguridad. El tiempo necesario para instalar el servidor y restaurar la copia de seguridad depende del tamaño de la copia y la complejidad del servidor. Pero es posible hacer todo fácilmente en cuestión de horas.

Soluciones de recuperación ante desastres de terceros

Existen muchas soluciones de terceros disponibles que proporcionan RPO de 15 minutos o más, y que sus equipos de TI probablemente ya utilicen. Aunque las soluciones de recuperación de copias de seguridad y recuperación ante desastres de terceros no se tratan en este informe, cuando se utiliza con dichos sistemas de recuperación de copias de seguridad, Tableau Server debe recibir el mismo tratamiento que cualquier otra aplicación esencial en el sistema y la tecnología de recuperación de copias de seguridad de terceros debe configurarse como corresponde.

Capacidades de recuperación ante desastres de Tableau

En el caso de que no pueda utilizar tecnologías de recuperación ante desastres de terceros, Tableau Server incluye una capacidad nativa incorporada de recuperación y copia de seguridad. Es posible crear copias de seguridad de Tableau Server mientras está en ejecución, por lo que puede obtener copias de seguridad

dinámicas y completas. A su vez, estas copias de seguridad dinámicas son tan inteligentes como Tableau Server y solo abarcan los archivos de extracción, las bases de datos de repositorio y la configuración correspondiente. También puede hacer copias de seguridad mientras Tableau Server no esté en ejecución (por ejemplo, durante un apagado planeado).

Puede optar por realizar una copia de seguridad de su servidor cada 24 horas si su RPO es de 24 horas, o bien cambiar el programa de copia de seguridad según su RPO. El tiempo de recuperación dependerá de si el error es de software o de hardware, además de su capacidad de proporcionar hardware nuevo en las instalaciones. Si dispone de otro nodo en espera con la misma versión de Tableau Server instalada y configurada, básicamente podrá restaurar Tableau Server a partir de la copia de seguridad mediante una operación rápida con comandos “admin” en la línea de comando.

Para la mayoría de las empresas, en las políticas de recuperación ante desastres también se imponen requisitos de distancias entre el centro de datos en el cual se ejecuta la copia principal del servidor y el centro de datos que representaría el objetivo de recuperación (p. ej., en Estados Unidos los sitios de recuperación de datos deben hallarse al menos a 250 millas). En estas situaciones, se puede considerar a Tableau como cualquier otra aplicación esencial del negocio en circunstancias de recuperación ante desastres. Aproveche las capacidades de recuperación y copia de seguridad incorporadas en Tableau para satisfacer sus necesidades de recuperación ante desastres.

Más allá de la alta disponibilidad

A fin de lograr una alta disponibilidad, las funcionalidades incorporadas en Tableau Server no son suficientes para proporcionar análisis de autoservicio a gran escala. Se pueden producir errores fuera del alcance del software de Tableau Server, como fallos en el hardware o en la red. Para lograr una alta disponibilidad mejorada, implemente Tableau Server en una infraestructura que sea resistente a estos tipos de errores, como la virtualización del hardware en nubes públicas o privadas. Algunos proveedores, como VMware vMotion, incluso ofrecen capacidades de migración en tiempo real.

Para lograr la alta disponibilidad verdadera, es necesario comprender las necesidades de los usuarios y seguir las prácticas recomendadas y los procesos para satisfacer dichas necesidades. Si bien las funcionalidades eficaces de Tableau facilitan la alta disponibilidad, no deberían reemplazar las prácticas recomendadas, como realizar mantenimiento habitualmente y planear copias de seguridad. Tableau es consciente de la importancia de las aplicaciones de inteligencia de negocios a la hora de tomar decisiones en su empresa. Tableau Server está preparado para ayudar a responder las preguntas de análisis más cruciales de su organización.

Acerca de Tableau

Tableau ayuda a las personas a transformar los datos en información útil para generar un impacto positivo. Conéctese con facilidad a datos almacenados en cualquier formato y lugar. Haga, rápidamente, análisis ad hoc que revelen oportunidades ocultas. Arrastre y suelte para crear dashboards interactivos con análisis visuales avanzados. Después, compártalos con toda su organización y permita que sus compañeros de equipo exploren los datos por sí mismos. Multinacionales, empresas pequeñas y emergentes... Todo el mundo usa la plataforma de análisis de Tableau para ver y comprender sus datos.

Recursos

[Tableau para la empresa: descripción general de TI](#)

[Tableau Server Scalability: A Technical Deployment Guide for Server Administrators \(Escalabilidad de Tableau Server: guía de implementación técnica para administradores\)](#)

[Guía de administrador de Tableau Server](#)

[Tableau on Amazon Web Services \(Tableau en Amazon Web Services\)](#)

