



Mantenimiento y mantenimiento

BeeGFS on NetApp with E-Series Storage

NetApp
June 18, 2024

Tabla de contenidos

- Mantenimiento y mantenimiento 1
 - Servicios de conmutación por error y conmutación tras recuperación 1
 - Coloque el clúster en modo de mantenimiento 3
 - Detenga e inicie el clúster 4
 - Sustituya los nodos de archivo 5
 - Expanda o reduzca el clúster 6

Mantenimiento y mantenimiento

Servicios de conmutación por error y conmutación tras recuperación

Desplazamiento de servicios BeeGFS entre nodos del clúster.

Descripción general

Los servicios BeeGFS pueden realizar una conmutación por error entre los nodos del clúster para garantizar que los clientes puedan continuar accediendo al sistema de archivos si un nodo experimenta un error o necesita realizar tareas de mantenimiento planificadas. En esta sección se describen distintas formas en las que los administradores pueden recuperar el clúster después de un fallo o mover manualmente servicios entre nodos.

Pasos

Conmutación por error y conmutación por recuperación

Conmutación al respaldo (planificada)

En general, cuando necesite desconectar un solo nodo de archivos para realizar el mantenimiento, querrá mover (o drenar) todos los servicios de BeeGFS de ese nodo. Esto se puede lograr poniendo el nodo en espera en primer lugar:

```
pcs node standby <HOSTNAME>
```

Después de verificar utilizando `pcs status` todos los recursos se han reiniciado en el nodo de archivos alternativo, puede apagar o realizar otros cambios en el nodo según sea necesario.

Conmutación tras recuperación (después de una conmutación al respaldo planificada)

Cuando esté listo para restaurar los servicios BeeGFS en el nodo preferido, ejecute primero `pcs status` Y verifique en la "Lista de nodos" el estado es en espera. Si el nodo se reinició, aparecerá sin conexión hasta que los servicios del clúster estén en línea:

```
pcs cluster start <HOSTNAME>
```

Una vez que el nodo esté en línea, salga del modo de espera con:

```
pcs cluster node unstandby <HOSTNAME>
```

Por último, reubique todos los servicios de BeeGFS en sus nodos preferidos con:

```
pcs resource relocate run
```

Conmutación tras recuperación (después de una conmutación al respaldo no planificada)

Si un nodo experimenta un fallo de hardware o de otro tipo, el clúster de alta disponibilidad debería reaccionar automáticamente y mover sus servicios a un nodo en buen estado, lo que proporciona tiempo para que los administradores tomen acciones correctivas. Antes de proceder, consulte "[resolución de problemas](#)" de la sección para determinar la causa de la conmutación por error y resolver cualquier problema pendiente. Una vez que el nodo se vuelve a encender y en buen estado, puede continuar con la conmutación tras recuperación.

Cuando un nodo se arranca tras un reinicio no planificado (o planificado), los servicios de clúster no se establecen para iniciarse automáticamente, por lo que primero tendrá que conectar el nodo con:

```
pcs cluster start <HOSTNAME>
```

A continuación, borre los errores de los recursos y restablezca el historial de cercas del nodo:

```
pcs resource cleanup node=<HOSTNAME>
pcs stonith history cleanup <HOSTNAME>
```

Verifique en `pcs status` el nodo está en línea y en buen estado. De forma predeterminada, los servicios de BeeGFS no se podrán recuperar automáticamente para evitar que los recursos vuelvan a un nodo que no esté en buenas estado. Cuando esté listo, devuelva todos los recursos del clúster a los nodos preferidos con:

```
pcs resource relocate run
```

Mover servicios de BeeGFS individuales a nodos de archivo alternativos

Mueva permanentemente un servicio BeeGFS a un nuevo nodo de archivo

Si desea cambiar de forma permanente el nodo de archivo preferido de un servicio BeeGFS individual, ajuste el inventario de Ansible para ver primero el nodo preferido y volver a ejecutar el libro de estrategia de Ansible.

Por ejemplo, en esta muestra `inventory.yml` File, `ictad22h01` es el nodo de archivo preferido para ejecutar el servicio de gestión BeeGFS:

```
  mgmt:
    hosts:
      ictad22h01:
      ictad22h02:
```

Si se invierte el pedido, se preferirían los servicios de gestión en `ictad22h02`:

```
mgmt:
  hosts:
    ictad22h02:
    ictad22h01:
```

Mueva temporalmente un servicio BeeGFS a otro nodo de archivo

Generalmente, si un nodo está en proceso de mantenimiento, deberá utilizar los [pasos de conmutación por error y conmutación por recuperación](#failover-and-failback) para mover todos los servicios fuera de ese nodo.

Si por algún motivo necesita mover un servicio individual a un nodo de archivo diferente ejecutado:

```
pcs resource move <SERVICE>-monitor <HOSTNAME>
```



No especifique recursos individuales ni el grupo de recursos. Especifique siempre el nombre del monitor para el servicio BeeGFS que desea reubicar. Por ejemplo, para mover el servicio de gestión de BeeGFS a ictad22h02, ejecute: `pcs resource move mgmt-monitor ictad22h02`. Este proceso se puede repetir para mover uno o varios servicios de sus nodos preferidos. Verifique utilizando `pcs status` los servicios se reubicaron o se iniciaron en el nuevo nodo.

Para devolver un servicio BeeGFS a su nodo preferido, borre primero las restricciones de recursos temporales (repita este paso según sea necesario para varios servicios):

```
pcs resource clear <SERVICE>-monitor
```

A continuación, cuando esté listo para mover realmente los servicios de nuevo a sus nodos preferidos ejecutar:

```
pcs resource relocate run
```

Nota este comando reubicará los servicios que ya no tengan restricciones temporales de recursos que no estén en sus nodos preferidos.

Coloque el clúster en modo de mantenimiento

Evite que el clúster de alta disponibilidad reaccione accidentalmente a los cambios previstos del entorno.

Descripción general

Si pone el clúster en el modo de mantenimiento, se deshabilita toda la supervisión de recursos y se impide que Pacemaker mueva o gestione los recursos del clúster de algún otro modo. Todos los recursos permanecerán en ejecución en sus nodos originales, independientemente de que haya una condición de fallo

temporal que impida que se pueda acceder a ellos. Los escenarios en los que esto es recomendable/útil incluyen:

- Mantenimiento de red que puede interrumpir temporalmente las conexiones entre nodos de archivo y servicios BeeGFS.
- Actualizaciones de nodos de bloques.
- Actualizaciones del sistema operativo, el kernel u otros paquetes del nodo de archivos.

Por lo general, el único motivo para poner manualmente el clúster en modo de mantenimiento es impedir que este reaccione a cambios externos en el entorno. Si un nodo individual del clúster requiere reparación física no utilice modo de mantenimiento y simplemente coloque ese nodo en espera tras el procedimiento anterior. Tenga en cuenta que el nuevo enrutamiento de Ansible pondrá automáticamente el clúster en modo de mantenimiento para facilitar la mayoría de las tareas de mantenimiento de software, incluidas las actualizaciones y los cambios de configuración.

Pasos

Para comprobar si el clúster se encuentra en modo de mantenimiento ejecutar:

```
pcs property show maintenance-mode
```

Esto devolverá el valor false cuando el clúster esté funcionando normalmente. Para habilitar la ejecución del modo de mantenimiento:

```
pcs property set maintenance-mode=true
```

Puede verificar ejecutando el estado del pc y asegurando que todos los recursos muestren "(no administrado)". Para desconectar el clúster del modo de mantenimiento ejecute:

```
pcs property set maintenance-mode=false
```

Detenga e inicie el clúster

Detener e iniciar correctamente el clúster de alta disponibilidad.

Descripción general

En esta sección se describe cómo apagar y reiniciar correctamente el clúster BeeGFS. Algunos ejemplos de casos en los que esto puede ser necesario son el mantenimiento eléctrico o la migración entre centros de datos o racks.

Pasos

Si por algún motivo necesita detener todo el clúster BeeGFS y apagar todos los servicios que se ejecutan:

```
pcs cluster stop --all
```

También es posible detener el clúster en nodos individuales (lo que conmuta automáticamente al nodo de respaldo en otro nodo), aunque se recomienda poner primero el nodo en espera (consulte la "[conmutación al respaldo](#)" sección):

```
pcs cluster stop <HOSTNAME>
```

Para iniciar los servicios y recursos del clúster en todos los nodos ejecutados:

```
pcs cluster start --all
```

O inicie servicios en un nodo específico con:

```
pcs cluster start <HOSTNAME>
```

En este momento, corre `pcs status` Compruebe que el clúster y los servicios BeeGFS se inicien en todos los nodos y que los servicios se estén ejecutando en los nodos que espera.



Según el tamaño del clúster, puede tardar en detenerse de algún momento (segundos a minutos) para que todo el clúster se detenga o se muestre iniciado en `pcs status`. Si `pcs cluster <COMMAND>` Se bloquea durante más de cinco minutos antes de ejecutar "Ctrl+C" para cancelar el comando, iniciar sesión en cada nodo del clúster y utilizar `pcs status` Para ver si los servicios de clúster (Corosync/Pacemaker) aún se están ejecutando en ese nodo. Desde cualquier nodo en el que el clúster siga estando activo, puede comprobar qué recursos están bloqueando el clúster. Solucione manualmente el problema y el comando debería estar completo o se puede volver a ejecutar para detener el resto de servicios.

Sustituya los nodos de archivo

Reemplazar un nodo de archivo si el servidor original está defectuoso.

Descripción general

Esta es una descripción general de los pasos necesarios para reemplazar un nodo de archivo del clúster. Estos pasos suponen un error en el nodo de archivo debido a un problema de hardware y se reemplazaron por un nuevo nodo de archivo idéntico.

Pasos:

1. Sustituya físicamente el nodo de archivo y restaure todo el cableado al nodo de bloque y a la red de almacenamiento.
2. Vuelva a instalar el sistema operativo en el nodo de archivos, incluida la adición de suscripciones a Red Hat.

3. Configure la gestión y las redes de BMC en el nodo de archivo.
4. Actualice el inventario de Ansible si el nombre de host, la IP, las asignaciones de interfaz de PCIe a lógica o cualquier otro cambio relacionado con el nodo de archivo nuevo. Generalmente, esto no es necesario si el nodo se reemplazó con un hardware de servidor idéntico y si utiliza la configuración de red original.
 - a. Por ejemplo, si el nombre de host ha cambiado, cree (o cambie de nombre) el archivo de inventario del nodo (`host_vars/<NEW_NODE>.yaml`) Luego en el archivo de inventario de Ansible (`inventory.yaml`), reemplace el nombre del nodo antiguo por el nuevo nombre del nodo:

```
all:
  ...
  children:
    ha_cluster:
      children:
        mgmt:
          hosts:
            node_h1_new:  # Replaced "node_h1" with "node_h1_new"
            node_h2:
```

5. De uno de los otros nodos del clúster, quite el nodo antiguo: `pcs cluster node remove <HOSTNAME>`.



NO CONTINÚE ANTES DE EJECUTAR ESTE PASO.

6. En el nodo de control Ansible:
 - a. Quite la clave SSH antigua con:

```
`ssh-keygen -R <HOSTNAME_OR_IP>`
```

- b. Configure SSH sin contraseña al nodo de sustitución con:

```
ssh-copy-id <USER>@<HOSTNAME_OR_IP>
```

7. Vuelva a ejecutar el libro de estrategia de Ansible para configurar el nodo y añadirlo al clúster:

```
ansible-playbook -i <inventory>.yaml <playbook>.yaml
```

8. En este momento, corre `pcs status` y compruebe que el nodo sustituido ahora aparece y ejecuta servicios.

Expanda o reduzca el clúster

Añada o quite bloques de creación del clúster.

Descripción general

En esta sección se documentan varias consideraciones y opciones para ajustar el tamaño del clúster de alta disponibilidad de BeeGFS. Normalmente, el tamaño del clúster se ajusta agregando o quitando elementos básicos, que normalmente son dos nodos de archivo configurados como un par de alta disponibilidad. También es posible agregar o quitar nodos de archivos individuales (u otros tipos de nodos de clústeres) si es necesario.

Añadir una elemento básico al clúster

Consideraciones

El crecimiento del clúster mediante la adición de elementos básicos adicionales es un proceso sencillo. Antes de empezar a tener en cuenta las restricciones en torno al número mínimo y máximo de nodos de clúster en cada clúster de alta disponibilidad, y determinar si debe añadir nodos al clúster de alta disponibilidad existente o crear un nuevo clúster de alta disponibilidad. Normalmente, cada bloque de creación consta de dos nodos de archivo, pero tres nodos son el número mínimo de nodos por clúster (para establecer quórum), y diez son el máximo recomendado (probado). En situaciones avanzadas es posible añadir un solo nodo "tiebreaker" que no ejecute ningún servicio BeeGFS al poner en marcha un clúster de dos nodos. Póngase en contacto con el servicio de soporte de NetApp si está considerando realizar dicha implementación.

Tenga en cuenta estas restricciones y cualquier crecimiento futuro del clúster anticipado al decidir cómo ampliar el clúster. Por ejemplo, si tiene un clúster de seis nodos y necesita añadir cuatro nodos más, se recomienda simplemente iniciar un nuevo clúster de alta disponibilidad.



Recuerde que un solo sistema de archivos BeeGFS puede consistir en varios clústeres de alta disponibilidad independientes. Esto permite que los sistemas de archivos sigan escalando entre los límites de hardware recomendado y los componentes de clúster de alta disponibilidad subyacentes.

Pasos

A la hora de añadir un elemento de creación al clúster, deberá crear el `host_vars` Archivos para cada uno de los nuevos nodos de archivos y nodos de bloques (cabines E-Series). Los nombres de estos hosts deben agregarse al inventario, junto con los nuevos recursos que se van a crear. El correspondiente `group_vars` será necesario crear archivos para cada recurso nuevo. Consulte "[Utilizar arquitecturas personalizadas](#)" para obtener más información.

Una vez creados los archivos correctos, todo lo que se necesita es volver a ejecutar la automatización mediante el comando:

```
ansible-playbook -i <inventory>.yaml <playbook>.yaml
```

Eliminación de un elemento básico del clúster

Hay una serie de consideraciones que tener en cuenta cuando se necesita retirar un bloque de construcción, por ejemplo:

- ¿Qué servicios BeeGFS se están ejecutando en este bloque de creación?
- ¿Solo se retiran los nodos de archivos y los nodos de bloque deben adjuntarse a nuevos nodos de archivos?

- Si se retira todo el elemento básico, ¿deben moverse los datos a un nuevo elemento básico, dispersarse en nodos existentes en el clúster o moverse a un nuevo sistema de archivos BeeGFS u otro sistema de almacenamiento?
- ¿Puede suceder esto durante una interrupción o se debe realizar de forma no disruptiva?
- ¿El elemento básico se está utilizando de forma activa o contiene principalmente datos que ya no están activos?

Debido a la diversidad de puntos de partida posibles y a los estados finales deseados, póngase en contacto con el soporte de NetApp para que podamos identificar y ayudar a implementar la mejor estrategia en función de su entorno y sus requisitos.

Información de copyright

Copyright © 2024 NetApp, Inc. Todos los derechos reservados. Imprimido en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPTIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en <http://www.netapp.com/TM> son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.