



Entretien et maintenance

BeeGFS on NetApp with E-Series Storage

NetApp
August 23, 2024

Sommaire

- Entretien et maintenance 1
 - Services de basculement/rétablissement 1
 - Placer le cluster en mode maintenance 3
 - Arrêtez et démarrez le cluster 4
 - Remplacer les nœuds de fichiers 5
 - Développez ou réduisez le cluster 6

Entretien et maintenance

Services de basculement/rétablissement

Déplacement des services BeeGFS entre les nœuds du cluster.

Présentation

Les services BeeGFS peuvent basculer entre les nœuds du cluster pour s'assurer que les clients sont en mesure de continuer à accéder au système de fichiers en cas de défaillance d'un nœud ou si vous devez effectuer une maintenance planifiée. Cette section décrit différentes méthodes permettant aux administrateurs d'effectuer une réparation sur le cluster après une reprise d'activité ou de déplacer manuellement les services entre les nœuds.

Étapes

Basculement et rétablissement

Basculement (planifié)

De manière générale, lorsque vous devez mettre un nœud de fichier unique hors ligne pour les opérations de maintenance, vous devez déplacer (ou vidanger) tous les services BeeGFS depuis ce nœud. Pour ce faire, le nœud peut d'abord être en veille :

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

Après vérification de l'utilisation `pcs status` toutes les ressources ont été redémarrées sur le nœud de fichier secondaire, vous pouvez arrêter ou apporter d'autres modifications au nœud si nécessaire.

Restauration (après un basculement planifié)

Lorsque vous êtes prêt à restaurer les services BeeGFS sur le nœud préféré s'exécutent d'abord `pcs status` Et vérifiez dans la « liste de nœuds » que l'état est en veille. Si le nœud a été redémarré, il s'affiche hors ligne jusqu'à ce que vous mettent les services du cluster en ligne :

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

Une fois le nœud mis en ligne hors veille, grâce à :

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

Enfin, transférez tous les services BeeGFS vers leurs nœuds préférés avec :

```
pcs resource relocate run
```

Retour arrière (après basculement non planifié)

Si un nœud présente un défaut matériel ou autre, le cluster haute disponibilité doit réagir automatiquement et déplacer ses services vers un nœud sain, ce qui permet aux administrateurs de prendre des mesures correctives. Avant de continuer, reportez-vous à "[dépannage](#)" la section pour déterminer la cause du basculement et résoudre tout problème en suspens. Une fois le nœud mis sous tension et en bon état, vous pouvez continuer à le restaurer.

Lorsqu'un nœud démarre après un redémarrage non planifié (ou planifié), les services de cluster ne sont pas configurés pour démarrer automatiquement. Vous devez donc mettre le nœud en ligne avec :

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

Ensuite, nettoyez toute défaillance de ressource et réinitialisez l'historique d'escrime du nœud :

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

Vérifier dans `pcs status` le nœud est en ligne et fonctionne correctement. Par défaut, les services BeeGFS ne sont pas automatiquement rebasculer afin d'éviter tout déplacement accidentel des ressources vers un nœud malsain. Une fois que vous êtes prêt à renvoyer toutes les ressources du cluster à leurs nœuds préférés avec :

```
pcs resource relocate run
```

Déplacement de services BeeGFS individuels vers d'autres nœuds de fichiers

Déplacer définitivement un service BeeGFS vers un nouveau nœud de fichier

Si vous souhaitez modifier de manière permanente le nœud de fichier favori pour un service BeeGFS, ajustez l'inventaire Ansible de sorte que le nœud préféré soit répertorié en premier et exécutez à nouveau le PlayBook Ansible.

Par exemple, dans cet exemple de `inventory.yml` fichier, `beegfs_01` est le nœud de fichiers préféré pour exécuter le service de gestion BeeGFS :

```
  mgmt:
    hosts:
      beegfs_01:
      beegfs_02:
```

Inverser l'ordre ferait que les services de gestion seraient préférés sur `beegfs_02`:

```
mgmt:
  hosts:
    beegfs_02:
    beegfs_01:
```

Déplacer temporairement un service BeeGFS vers un autre nœud de fichier

De manière générale, si un nœud est en cours de maintenance, il convient d'utiliser [les étapes de basculement et de retour arrière](#le-basculement-et-la-restauration) pour déplacer tous les services hors de ce nœud.

Si vous devez déplacer un service individuel vers un autre nœud de fichiers :

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



Ne spécifiez pas les ressources individuelles ou le groupe de ressources. Spécifiez toujours le nom du moniteur pour le service BeeGFS que vous souhaitez déplacer. Par exemple, pour déplacer le service de gestion BeeGFS vers `beegfs_02`, exécutez : `pcs resource move mgmt-monitor beegfs_02`. Ce processus peut être répété afin de déplacer un ou plusieurs services hors de leurs nœuds préférés. Vérifiez à l'aide des `pcs status services` qui ont été déplacés/démarrés sur le nouveau nœud.

Pour déplacer un service BeeGFS vers son nœud préféré, effacez d'abord les contraintes de ressources temporaires (en répétant cette étape comme nécessaire pour plusieurs services) :

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

Ensuite, une fois prêt à rapatrier les services sur les nœuds de leur choix :

```
pcs resource relocate run
```

Notez que cette commande permet de transférer tous les services qui ne disposent plus de contraintes temporaires en termes de ressources, situés sur les nœuds de leur choix.

Placer le cluster en mode maintenance

Empêcher le cluster de haute disponibilité de réagir accidentellement aux changements prévus dans l'environnement.

Présentation

Le fait de mettre le cluster en mode maintenance désactive toute la surveillance des ressources et empêche Pacemaker de déplacer ou de gérer des ressources dans le cluster. Toutes les ressources restent exécutées sur leurs nœuds d'origine, peu importe la condition de panne temporaire qui empêcherait leur accès. Voici quelques scénarios recommandés/utiles :

- Maintenance du réseau pouvant interrompre temporairement les connexions entre les nœuds de fichiers et les services BeeGFS.
- Mises à niveau des nœuds de blocs.
- Mises à jour du système d'exploitation de nœud de fichiers, du noyau ou d'autres modules.

En général, la seule raison de placer manuellement le cluster en mode de maintenance est d'éviter que le système ne réagisse à des modifications externes de l'environnement. Si un nœud individuel du cluster nécessite une réparation physique, n'utilisez pas le mode de maintenance et placez simplement ce nœud en veille après la procédure ci-dessus. Notez que le changement d'Ansible place automatiquement le cluster en mode de maintenance pour faciliter la plupart des opérations de maintenance logicielle, y compris les mises à niveau et les modifications de configuration.

Étapes

Pour vérifier si le cluster est en mode maintenance, exécutez :

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

Cela renvoie la valeur FALSE lorsque le cluster fonctionne normalement. Pour activer le mode maintenance, exécutez :

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

Vous pouvez vérifier en exécutant l'état pcs et en vous assurant que toutes les ressources affichent « (non géré) ». Pour mettre le cluster hors mode maintenance, exécutez :

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

Arrêtez et démarrez le cluster

Arrêt et démarrage du cluster HA avec élégance.

Présentation

Cette section décrit comment arrêter et redémarrer le cluster BeeGFS. Par exemple, la maintenance électrique ou la migration d'un data Center à l'autre ou d'un rack peut être nécessaire.

Étapes

Si, pour une raison quelconque, vous devez arrêter tout le cluster BeeGFS et arrêter tous les services exécutent :

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

Il est également possible d'arrêter le cluster sur des nœuds individuels (qui basculeront automatiquement les

services vers un autre nœud), bien qu'il soit d'abord recommandé de mettre le nœud en veille (voir la "basculement" section) :

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

Pour démarrer les ressources et les services du cluster sur tous les nœuds, exécutez :

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

Ou démarrer les services sur un nœud spécifique avec :

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

À ce stade, exécuter `pcs status` Et vérifiez que le cluster et les services BeeGFS démarrent sur tous les nœuds et que les services sont exécutés sur les nœuds que vous attendez.



Selon la taille du cluster, l'arrêt de l'ensemble du cluster peut prendre quelques secondes ou minutes, voire d'autres opérations `pcs status`. Si `pcs cluster <COMMAND>` Cesse de répondre pendant plus de cinq minutes, avant d'exécuter « Ctrl+C » pour annuler la commande, connectez-vous à chaque nœud du cluster et utilisez `pcs status` Pour vérifier si les services de cluster (Corosync/Pacemaker) sont toujours en cours d'exécution sur ce nœud. A partir de n'importe quel nœud où le cluster est toujours actif, vous pouvez vérifier les ressources qui bloquent le cluster. Résoudre manuellement le problème et la commande doit être terminée ou réexécutée pour arrêter les services restants.

Remplacer les nœuds de fichiers

Remplacement d'un noeud de fichier si le serveur d'origine est défectueux.

Présentation

Voici un aperçu des étapes nécessaires au remplacement d'un noeud de fichier dans le cluster. Ces étapes présupposent que le nœud de fichier a échoué en raison d'un problème matériel et a été remplacé par un nouveau nœud de fichier identique.

Étapes :

1. Remplacez physiquement le nœud de fichiers et restaurez tout le câblage vers le nœud de bloc et le réseau de stockage.
2. Réinstallez le système d'exploitation sur le nœud de fichier, y compris l'ajout d'abonnements Red Hat.
3. Configurez la mise en réseau BMC et la gestion sur le nœud de fichiers.
4. Mettez à jour l'inventaire Ansible si le nom d'hôte, l'IP, les mappages de l'interface PCIe vers l'interface logique ou tout autre élément modifié concernant le nouveau nœud de fichier. En général, cette opération n'est pas nécessaire si le nœud a été remplacé par le même matériel serveur et que vous utilisez la configuration réseau d'origine.

- a. Par exemple, si le nom d'hôte a changé, créez (ou renommez) le fichier d'inventaire du nœud (`host_vars/<NEW_NODE>.yaml`) Puis dans le fichier d'inventaire Ansible (`inventory.yaml`), remplacez le nom de l'ancien nœud par le nouveau nom de nœud :

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

5. Depuis un des autres nœuds du cluster, supprimer l'ancien nœud : `pcs cluster node remove <HOSTNAME>`.



NE PAS POURSUIVRE AVANT D'EXÉCUTER CETTE ÉTAPE.

6. Sur le nœud de contrôle Ansible :
 - a. Supprimez l'ancienne clé SSH avec :

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

- b. Configurez SSH sans mot de passe sur le nœud remplacer par :

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

7. Exécutez à nouveau le PlayBook Ansible pour configurer le nœud et l'ajouter au cluster :

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

8. A ce stade, exécuter `pcs status` et vérifiez que le nœud remplacé est maintenant répertorié et que les services sont en cours d'exécution.

Développez ou réduisez le cluster

Ajouter ou supprimer des éléments de base du cluster.

Présentation

Cette section présente divers éléments à prendre en compte et diverses options pour ajuster la taille de votre cluster BeeGFS HA. La taille du cluster est généralement ajustée en ajoutant ou en supprimant des éléments de base, qui sont généralement deux nœuds de fichiers configurés comme une paire haute disponibilité. Il est

également possible d'ajouter ou de supprimer des nœuds de fichiers individuels (ou d'autres types de nœuds de cluster) si nécessaire.

Ajout d'un module au cluster

Considérations

Le développement du cluster par l'ajout d'éléments de base supplémentaires est un processus simple. Avant de commencer, les restrictions s'imposent concernant le nombre minimal et maximal de nœuds de cluster dans chaque cluster haute disponibilité. Déterminer si vous devez ajouter des nœuds au cluster haute disponibilité existant ou créer un nouveau cluster haute disponibilité. En général, chaque élément de base se compose de deux nœuds de fichiers, mais trois nœuds représentent le nombre minimum de nœuds par cluster (pour établir le quorum) et dix est le nombre maximum recommandé (testé). Pour les scénarios avancés, il est possible d'ajouter un nœud « Tiebreaker » unique qui n'exécute aucun service BeeGFS lors du déploiement d'un cluster à deux nœuds. Si vous envisagez un tel déploiement, contactez le support NetApp.

Gardez à l'esprit ces restrictions et toute future croissance des clusters lorsque vous décidez d'étendre le cluster. Par exemple, si vous disposez d'un cluster à six nœuds et que vous devez en ajouter quatre autres, il est recommandé de simplement démarrer un nouveau cluster haute disponibilité.



N'oubliez pas qu'un seul système de fichiers BeeGFS peut consister en plusieurs clusters HA indépendants. Les systèmes de fichiers peuvent ainsi continuer à évoluer au-delà des limites recommandées/strictes des composants de cluster haute disponibilité sous-jacents.

Étapes

Lorsque vous ajoutez un élément de base à votre cluster, vous devez créer les `host_vars` fichiers pour chacun des nouveaux nœuds de fichiers et nœuds de blocs (baies E-Series). Les noms de ces hôtes doivent être ajoutés à l'inventaire, ainsi que les nouvelles ressources à créer. Les `group_vars` fichiers correspondants devront être créés pour chaque nouvelle ressource. Voir la "[utilisez des architectures personnalisées](#)" section pour plus de détails.

Une fois les fichiers corrects créés, il suffit de relancer l'automatisation à l'aide de la commande :

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

Retrait d'un module du cluster

Il y a plusieurs considérations à garder à l'esprit lorsque vous devez retirer un élément de construction, par exemple :

- Quels sont les services BeeGFS exécutés dans cet élément de base ?
- Les nœuds de fichiers ne sont-ils que ceux qui sont mis hors service et ceux qui doivent être associés à de nouveaux nœuds de fichiers ?
- Si l'ensemble du bloc de construction est retiré, les données doivent-elles être déplacées vers un nouveau bloc de construction, dispersées vers les nœuds existants du cluster ou déplacées vers un nouveau système de fichiers BeeGFS ou un autre système de stockage ?
- Cela peut-il avoir lieu en cas de panne ou doit-il être effectué sans interruption ?
- L'élément de base est-il activement utilisé ou contient-il principalement des données qui ne sont plus actives ?

Étant donné la diversité des points de départ et des États de terminaison, veuillez contacter le support NetApp afin que nous puissions identifier et vous aider à mettre en œuvre la meilleure stratégie en fonction de votre environnement et de vos besoins.

Informations sur le copyright

Copyright © 2024 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTEUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.