



TR-4956 : AWS FSx/EC2 中的自動化 PostgreSQL 高可用性部署與災難復原

NetApp database solutions

NetApp
August 18, 2025

目錄

TR-4956：AWS FSx/EC2 中的自動化 PostgreSQL 高可用性部署與災難復原	1
目的	1
對象	1
解決方案測試和驗證環境	1
架構	2
硬體和軟體組件	2
部署考慮的關鍵因素	3
解決方案部署	3
自動部署的先決條件	3
設定 hosts 文件	4
配置host_vars資料夾中的host_name.yml文件	5
配置vars資料夾中的全域fsx_vars.yml文件	6
PostgreSQL 部署和 HA/DR 設定	8
PostgreSQL 資料庫快照備份並複製到備用站點	9
故障轉移到災難復原備用站點	10
故障轉移測試後重新同步複製的資料庫卷	10
由於 EC2 計算執行個體故障，從主 EC2 資料庫伺服器故障轉移到備用 EC2 資料庫伺服器	10
在哪裡可以找到更多信息	10

TR-4956：AWS FSx/EC2 中的自動化 PostgreSQL 高可用性部署與災難復原

Allen Cao、Niyaz Mohamed，NetApp

該解決方案提供了基於 FSx ONTAP 儲存產品內建的 NetApp SnapMirror 技術和 AWS 中的 NetApp Ansible 自動化工具包的 PostgreSQL 資料庫部署和 HA/DR 設定、故障轉移、重新同步的概述和詳細資訊。

目的

PostgreSQL 是一個廣泛使用的開源資料庫，在十大最受歡迎資料庫引擎中排名第四。["資料庫引擎"](#)。一方面，PostgreSQL 因其免許可、開源模式而廣受歡迎，同時仍具有複雜的功能。另一方面，由於它是開源的，因此缺乏關於高可用性和災難復原（HA/DR）領域的生產級資料庫部署的詳細指導，尤其是在公有雲中。一般來說，建立一個具有熱備用、溫備用、流複製等功能的典型 PostgreSQL HA/DR 系統可能很困難。透過提升備用站點然後切換回主站點來測試 HA/DR 環境可能會對生產造成干擾。當讀取工作負載部署在串流熱備用系統上時，主資料庫上有充分記錄的效能問題。

在本文檔中，我們示範如何擺脫應用程式層級 PostgreSQL 串流 HA/DR 解決方案，並使用儲存層級複製建置基於 AWS FSx ONTAP 儲存和 EC2 運算執行個體的 PostgreSQL HA/DR 解決方案。與傳統的 PostgreSQL 應用程式級 HA/DR 流複製相比，該解決方案創建了一個更簡單、更可比系統，並提供了相同的結果。

該解決方案基於經過驗證的成熟 NetApp SnapMirror 儲存級複製技術構建，該技術可用於 AWS 原生 FSX ONTAP 雲端儲存，以實現 PostgreSQL HA/DR。使用 NetApp 解決方案團隊提供的自動化工具包可以輕鬆實現。它提供類似的功能，同時透過基於應用程式級流的 HA/DR 解決方案消除主站點的複雜性和效能拖累。該解決方案可以輕鬆部署和測試，而不會影響活動主站點。

此解決方案適用於以下用例：

- 在公用 AWS 雲端中為 PostgreSQL 提供生產級 HA/DR 部署
- 在公用 AWS 雲端中測試和驗證 PostgreSQL 工作負載
- 測試並驗證基於 NetApp SnapMirror 複製技術的 PostgreSQL HA/DR 策略

對象

此解決方案適用於以下人群：

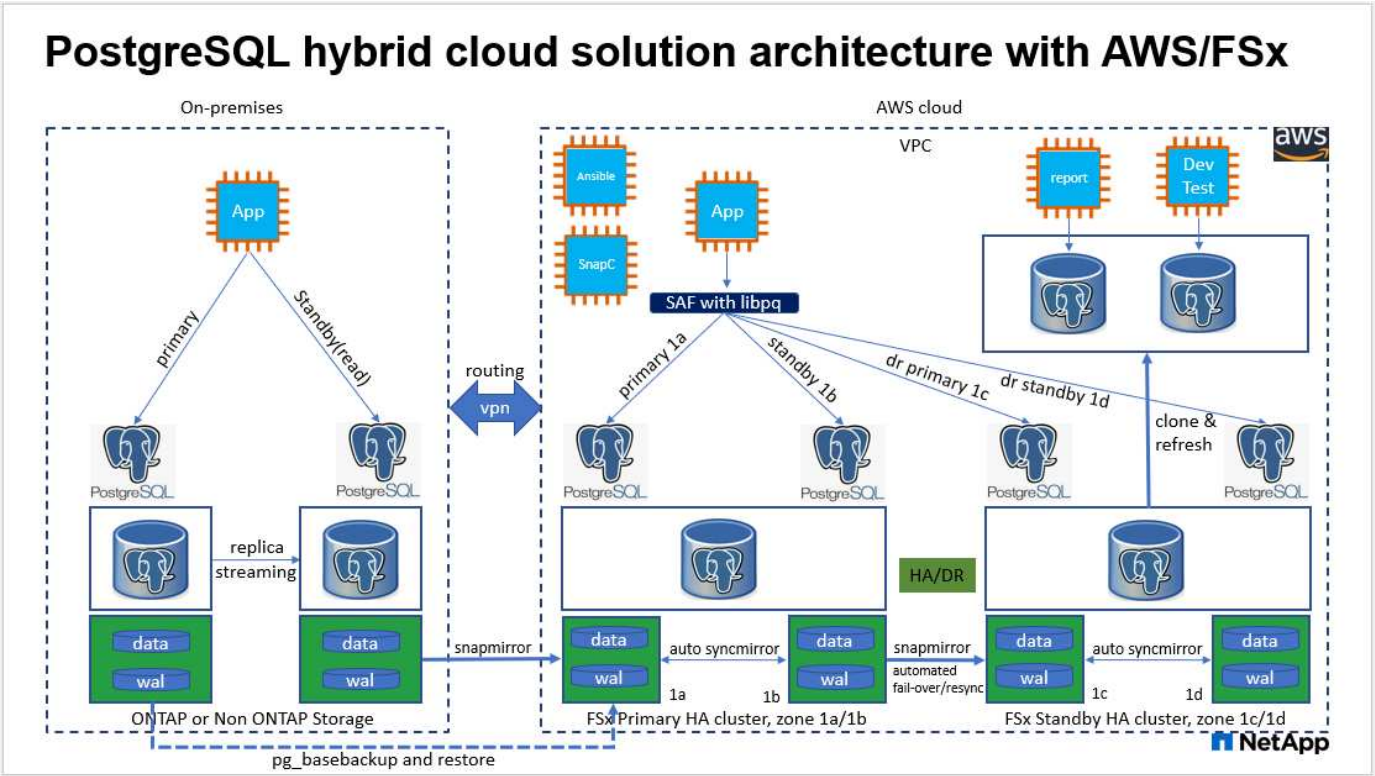
- 有興趣在公有 AWS 雲端中部署具有 HA/DR 的 PostgreSQL 的 DBA。
- 對在公用 AWS 雲端中測試 PostgreSQL 工作負載感興趣的資料庫解決方案架構師。
- 有興趣部署和管理部署到 AWS FSx 儲存的 PostgreSQL 實例的儲存管理員。
- 有興趣在 AWS FSx/EC2 中建立 PostgreSQL 環境的應用程式擁有者。

解決方案測試和驗證環境

此解決方案的測試和驗證是在可能與最終部署環境不符的 AWS FSx 和 EC2 環境中進行的。有關更多信息，請

參閱[部署考慮的關鍵因素]。

架構



硬體和軟體組件

硬體		
FSx ONTAP存儲	目前版本	同一 VPC 和可用區中的兩個 FSx HA 對作為主 HA 集群和備用 HA 集群
用於計算的 EC2 實例	t2.xlarge/4vCPU/16G	兩個 EC2 T2 xlarge 作為主運算執行個體和備用運算執行個體
Ansible 控制器	本地 Centos VM/4vCPU/8G	用於在本地或雲端託管 Ansible 自動化控制器的虛擬機
軟體		
紅帽Linux	RHEL-8.6.0_HVM-20220503-x86_64-2-Hourly2-GP2	部署 RedHat 訂閱進行測試
Centos Linux	CentOS Linux 版本 8.2.2004 (核心)	在本地實驗室中部署託管 Ansible 控制器
PostgreSQL	版本 14.5	自動化從 postgresql.ora yum repo 中提取最新可用的 PostgreSQL 版本
Ansible	版本 2.10.3	使用需求手冊安裝所需集合和函式庫的先決條件

部署考慮的關鍵因素

- *PostgreSQL 資料庫備份、還原和還原。* PostgreSQL 資料庫支援多種備份方法，例如使用 `pg_dump` 進行邏輯備份、使用 `pg_basebackup` 或較低層級的 OS 備份命令進行實體線上備份以及儲存層級一致的快照。此解決方案使用 NetApp 一致性群組快照在備用站點對 PostgreSQL 資料庫資料和 WAL 磁碟區進行備份、還原和還原。NetApp 一致性群組磁碟區快照在將 I/O 寫入儲存時對其進行排序，並保護資料庫資料檔案的完整性。
- EC2 運算執行個體。在這些測試和驗證中，我們使用 AWS EC2 `t2.xlarge` 執行個體類型作為 PostgreSQL 資料庫計算執行個體。NetApp 建議在部署中使用 M5 類型的 EC2 執行個體作為 PostgreSQL 的計算實例，因為它針對資料庫工作負載進行了最佳化。備用計算實例應始終與為 FSx HA 叢集部署的被動（備用）檔案系統部署在同一區域中。
- *FSx 儲存 HA 叢集單區域或多區域部署。* 在這些測試和驗證中，我們在單一 AWS 可用區中部署了一個 FSx HA 叢集。對於生產部署，NetApp 建議在兩個不同的可用區部署 FSx HA 對。如果主伺服器與備用伺服器之間需要特定的距離，則可以在不同區域設定災難復原備用 HA 對，以實現業務連續性。FSx HA 叢集始終在 HA 對中配置，該 HA 對在主動-被動檔案系統中同步鏡像，以提供儲存級冗餘。
- PostgreSQL 資料和日誌放置。典型的 PostgreSQL 部署共用相同的根目錄或磁碟區用於資料和日誌檔案。在我們的測試和驗證中，我們將 PostgreSQL 資料和日誌分成兩個單獨的磁碟區以提高效能。資料目錄中使用軟連結指向託管 PostgreSQL WAL 日誌和存檔 WAL 日誌的日誌目錄或磁碟區。
- *PostgreSQL 服務啟動延遲計時器。* 此解決方案使用 NFS 掛載磁碟區來儲存 PostgreSQL 資料庫檔案和 WAL 日誌檔案。在資料庫主機重新啟動期間，PostgreSQL 服務可能會在未安裝磁碟區的情況下嘗試啟動。這會導致資料庫服務啟動失敗。PostgreSQL 資料庫需要 10 到 15 秒的計時器延遲才能正確啟動。
- *RPO/RTO 確保業務連續性。* FSx 資料從主伺服器到備用伺服器的 DR 複製是基於 ASYNC，這意味著 RPO 取決於 Snapshot 備份和 SnapMirror 複製的頻率。更高的 Snapshot 副本和 SnapMirror 複製頻率可降低 RPO。因此，災難發生時的潛在資料遺失和增量儲存成本之間存在平衡。我們已經確定，Snapshot 副本和 SnapMirror 複製可以在低至 5 分鐘的間隔內實現 RPO，並且 PostgreSQL 通常可以在一分鐘內於 DR 備用站點恢復 RTO。
- *資料庫備份。* 將 PostgreSQL 資料庫從本機資料中心實作或移轉到 AWS FSx 儲存後，資料會在 FSx HA 對中自動同步鏡像以進行保護。一旦發生災難，資料將透過複製的備用站點得到進一步保護。對於長期備份保留或資料保護，NetApp 建議使用內建的 PostgreSQL `pg_basebackup` 公用程式來執行可移植到 S3 blob 儲存體的完整資料庫備份。

解決方案部署

按照下面概述的詳細說明，可以使用基於 NetApp Ansible 的自動化工具包自動完成此解決方案的部署。

1. 閱讀自動化工具包 README.md 中的說明["na_postgresql_aws_deploy_hadr"](#)。
2. 觀看以下影片示範。

自動化 PostgreSQL 部署與保護

1. 配置所需的參數文件(`hosts`，`host_vars/host_name.yml`，`fsx_vars.yml`) 透過在相關部分的範本中輸入使用者特定的參數。然後使用複製按鈕將檔案複製到 Ansible 控制器主機。

自動部署的先決條件

部署需要以下先決條件。

1. 已設定 AWS 帳戶，並在您的 AWS 帳戶內建立了必要的 VPC 和網路段。

2. 從 AWS EC2 控制台，您必須部署兩個 EC2 Linux 執行個體，一個作為主網站上的主 PostgreSQL DB 伺服器，另一個作為備用 DR 網站上的伺服器。為了在主 DR 站點和備用 DR 站點實現計算冗餘，請部署兩個額外的 EC2 Linux 執行個體作為備用 PostgreSQL DB 伺服器。有關環境設定的更多詳細信息，請參閱上一節中的架構圖。還請查看["Linux 實例使用者指南"](#)了解更多。
3. 從 AWS EC2 控制台部署兩個 FSx ONTAP 儲存 HA 叢集來託管 PostgreSQL 資料庫磁碟區。如果您不熟悉 FSx 儲存的部署，請參閱文檔["建立 FSx ONTAP 檔案系統"](#)以獲得逐步說明。
4. 建立一個 Centos Linux VM 來託管 Ansible 控制器。Ansible 控制器可以位於本機或 AWS 雲端。如果它位於本地，則必須具有與 VPC、EC2 Linux 實例和 FSx 儲存叢集的 SSH 連線。
5. 依照資源中的「在 RHEL/CentOS 上為 CLI 部署設定 Ansible 控制節點」部分所述設定 Ansible 控制器["NetApp 解決方案自動化入門"](#)。
6. 從公共 NetApp GitHub 網站複製自動化工具包的副本。

```
git clone https://github.com/NetApp-  
Automation/na_postgresql_aws_deploy_hadr.git
```

1. 從工具包根目錄執行先決條件劇本來安裝 Ansible 控制器所需的集合和函式庫。

```
ansible-playbook -i hosts requirements.yml
```

```
ansible-galaxy collection install -r collections/requirements.yml --force  
--force-with-deps
```

1. 檢索資料庫主機變數檔案所需的 EC2 FSx 執行個體參數 `host\_vars/\*` 和全域變數文件 `fsx\_vars.yml` 配置。

設定 hosts 文件

將主 FSx ONTAP 叢集管理 IP 和 EC2 執行個體主機名稱輸入到 hosts 檔案中。

```
# Primary FSx cluster management IP address  
[fsx_ontap]  
172.30.15.33
```

```
# Primary PostgreSQL DB server at primary site where database is  
initialized at deployment time  
[postgresql]  
psql_01p ansible_ssh_private_key_file=psql_01p.pem
```

```
# Primary PostgreSQL DB server at standby site where postgresql service is
installed but disabled at deployment
# Standby DB server at primary site, to setup this server comment out
other servers in [dr_postgresql]
# Standby DB server at standby site, to setup this server comment out
other servers in [dr_postgresql]
[dr_postgresql] --
psql_01s ansible_ssh_private_key_file=psql_01s.pem
#psql_01ps ansible_ssh_private_key_file=psql_01ps.pem
#psql_01ss ansible_ssh_private_key_file=psql_01ss.pem
```

配置**host_vars**資料夾中的**host_name.yml**文件

```
# Add your AWS EC2 instance IP address for the respective PostgreSQL
server host
ansible_host: "10.61.180.15"

# "{{groups.postgresql[0]}}" represents first PostgreSQL DB server as
defined in PostgreSQL hosts group [postgresql]. For concurrent multiple
PostgreSQL DB servers deployment, [0] will be incremented for each
additional DB server. For example, "{{groups.postgresql[1]}}" represents
DB server 2, "{{groups.postgresql[2]}}" represents DB server 3 ... As a
good practice and the default, two volumes are allocated to a PostgreSQL
DB server with corresponding /pgdata, /pglogs mount points, which store
PostgreSQL data, and PostgreSQL log files respectively. The number and
naming of DB volumes allocated to a DB server must match with what is
defined in global fsx_vars.yml file by src_db_vols, src_archivelog_vols
parameters, which dictates how many volumes are to be created for each DB
server. aggr_name is aggr1 by default. Do not change. lif address is the
NFS IP address for the SVM where PostgreSQL server is expected to mount
its database volumes. Primary site servers from primary SVM and standby
servers from standby SVM.
host_datastores_nfs:
  - {vol_name: "{{groups.postgresql[0]}}_pgdata", aggr_name: "aggr1", lif:
    "172.21.94.200", size: "100"}
  - {vol_name: "{{groups.postgresql[0]}}_pglogs", aggr_name: "aggr1", lif:
    "172.21.94.200", size: "100"}

# Add swap space to EC2 instance, that is equal to size of RAM up to 16G
max. Determine the number of blocks by dividing swap size in MB by 128.
swap_blocks: "128"

# Postgresql user configurable parameters
psql_port: "5432"
buffer_cache: "8192MB"
archive_mode: "on"
max_wal_size: "5GB"
client_address: "172.30.15.0/24"
```

配置vars資料夾中的全域fsx_vars.yml文件

```
#####
##### PostgreSQL HADR global user configuration variables #####
##### Consolidate all variables from FSx, Linux, and postgresql #####
#####

#####
```

```

### Ontap env specific config variables ###
#####

#####
#####
# Variables for SnapMirror Peering
#####
#####

#Passphrase for cluster peering authentication
passphrase: "xxxxxxx"

#Please enter destination or standby FSx cluster name
dst_cluster_name: "FsxId0cf8e0bccb14805e8"

#Please enter destination or standby FSx cluster management IP
dst_cluster_ip: "172.30.15.90"

#Please enter destination or standby FSx cluster inter-cluster IP
dst_inter_ip: "172.30.15.13"

#Please enter destination or standby SVM name to create mirror
relationship
dst_vserver: "dr"

#Please enter destination or standby SVM management IP
dst_vserver_mgmt_lif: "172.30.15.88"

#Please enter destination or standby SVM NFS lif
dst_nfs_lif: "172.30.15.88"

#Please enter source or primary FSx cluster name
src_cluster_name: "FsxId0cf8e0bccb14805e8"

#Please enter source or primary FSx cluster management IP
src_cluster_ip: "172.30.15.20"

#Please enter source or primary FSx cluster inter-cluster IP
src_inter_ip: "172.30.15.5"

#Please enter source or primary SVM name to create mirror relationship
src_vserver: "prod"

#Please enter source or primary SVM management IP
src_vserver_mgmt_lif: "172.30.15.115"

#####

```

```
#####
# Variable for PostgreSQL Volumes, lif - source or primary FSx NFS lif
address
#####
#####

src_db_vols:
  - {vol_name: "{{groups.postgresql[0]}}_pgdata", aggr_name: "aggr1", lif:
    "172.21.94.200", size: "100"}

src_archivelog_vols:
  - {vol_name: "{{groups.postgresql[0]}}_pglogs", aggr_name: "aggr1", lif:
    "172.21.94.200", size: "100"}

#Names of the Nodes in the ONTAP Cluster
nfs_export_policy: "default"

#####
#####
### Linux env specific config variables ###
#####
#####

#NFS Mount points for PostgreSQL DB volumes
mount_points:
  - "/pgdata"
  - "/pglogs"

#RedHat subscription username and password
redhat_sub_username: "xxxxx"
redhat_sub_password: "xxxxx"

#####
### DB env specific install and config variables ###
#####
#The latest version of PostgreSQL RPM is pulled/installed and config file
is deployed from a preconfigured template
#Recovery type and point: default as all logs and promote and leave all
PITR parameters blank
```

PostgreSQL 部署和 HA/DR 設定

以下任務在主 EC2 DB 伺服器主機上部署 PostgreSQL DB 伺服器服務並在主站點初始化資料庫。然後在備用站點設定備用主 EC2 DB 伺服器主機。最後，從主站點 FSx 叢集到備用站點 FSx 叢集建立 DB 磁碟區複製，以實現災難復原。

1. 在主 FSx 叢集上建立 DB 卷，並在主 EC2 執行個體主機上設定 postgresql。

```
ansible-playbook -i hosts postgresql_deploy.yml -u ec2-user --private-key psql_01p.pem -e @vars/fsx_vars.yml
```

2. 設定備用 DR EC2 執行個體主機。

```
ansible-playbook -i hosts postgresql_standby_setup.yml -u ec2-user --private-key psql_01s.pem -e @vars/fsx_vars.yml
```

3. 設定 FSx ONTAP 叢集對等和資料庫磁碟區複製。

```
ansible-playbook -i hosts fsx_replication_setup.yml -e @vars/fsx_vars.yml
```

4. 將前面的步驟合併為單步驟 PostgreSQL 部署和 HA/DR 設定。

```
ansible-playbook -i hosts postgresql_hadr_setup.yml -u ec2-user -e @vars/fsx_vars.yml
```

5. 若要在主網站或備用網站設定備用 PostgreSQL DB 主機，請註解掉主機檔案 [dr_postgresql] 部分中的所有其他伺服器，然後使用對應的目標主機（例如主網站上的 psql_01ps 或備用 EC2 運算執行個體）執行 postgresql_standby_setup.yml 劇本。確保主機參數檔如 `psql\_01ps.yml` 配置在 `host\_vars` 目錄。

```
[dr_postgresql] --
#psql_01s ansible_ssh_private_key_file=psql_01s.pem
psql_01ps ansible_ssh_private_key_file=psql_01ps.pem
#psql_01ss ansible_ssh_private_key_file=psql_01ss.pem
```

```
ansible-playbook -i hosts postgresql_standby_setup.yml -u ec2-user --private-key psql_01ps.pem -e @vars/fsx_vars.yml
```

PostgreSQL 資料庫快照備份並複製到備用站點

PostgreSQL 資料庫快照備份和複製到備用網站可以在 Ansible 控制器上以使用者定義的間隔進行控制和執行。我們已經驗證間隔可以低至 5 分鐘。因此，當主站點發生故障時，如果故障發生在下一次計劃快照備份之前，則可能會有 5 分鐘的資料遺失。

```
*/15 * * * * /home/admin/na_postgresql_aws_deploy_hadr/data_log_snap.sh
```

故障轉移到災難復原備用站點

為了將 PostgreSQL HA/DR 系統作為 DR 練習進行測試，請透過執行下列劇本在備用網站上的主備用 EC2 DB 執行個體上執行故障轉移和 PostgreSQL 資料庫復原。在實際的 DR 場景中，對實際故障轉移到 DR 站點執行相同的操作。

```
ansible-playbook -i hosts postgresql_failover.yml -u ec2-user --private-key psql_01s.pem -e @vars/fsx_vars.yml
```

故障轉移測試後重新同步複製的資料庫卷

故障轉移測試後執行重新同步以重新建立資料庫磁碟區 SnapMirror 複製。

```
ansible-playbook -i hosts postgresql_standby_resync.yml -u ec2-user --private-key psql_01s.pem -e @vars/fsx_vars.yml
```

由於 **EC2** 計算執行個體故障，從主 **EC2** 資料庫伺服器故障轉移到備用 **EC2** 資料庫伺服器

NetApp 建議執行手動故障轉移或使用可能需要許可證的成熟 OS 叢集軟體。

在哪裡可以找到更多信息

要了解有關本文檔中描述的信息的更多信息，請查看以下文檔和/或網站：

- Amazon FSx ONTAP

["https://aws.amazon.com/fsx/netapp-ontap/"](https://aws.amazon.com/fsx/netapp-ontap/)

- 亞馬遜 EC2

https://aws.amazon.com/pm/ec2/?trk=36c6da98-7b20-48fa-8225-4784bced9843&sc_channel=ps&s_kwcid=AL!4422!3!467723097970!e!!g!!aws%20ec2&ef_id=Cj0KCQiA54KfBhCKARIsAJzSrdqwQrghn6l71jiWzSeaT9Uh1-vY-VfhJixF-xnv5rWwn2S7RqZOTQ0aAh7eEALw_wcB:G:s&s_kwcid=AL!4422!3!467723097970!e!!g!!aws%20ec2

- NetApp 解決方案自動化

["介紹"](#)

版權資訊

Copyright © 2025 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。