

對機架間複製故障進行故障排除，錯誤代碼為 "424-Geo-replication Checksum Mismatch"

目錄

[簡介](#)

[必要條件](#)

[需求](#)

[採用元件](#)

[背景資訊](#)

[什麼是SMF中的Geo冗餘？](#)

[異地複製Pod](#)

[確定活動的Geo Pod和備用的Geo Pod](#)

[GR POD的功能](#)

[GR Pod跨站點複製ETCD和快取Pod資料](#)

[在ETCD中維護站點本地例項角色](#)

[監控本地站點狀態 \(POD狀態/BFD狀態 \)](#)

[站點角色](#)

[GR-Triggers](#)

[用於驗證機架上的GR例項角色的CLI](#)

[將角色從備用錯誤重置為備用的CLI](#)

[CLI to Switch Role from Standby to Standby錯誤](#)

[CLI將角色從備用狀態切換到主狀態](#)

[TCP連線終止](#)

[問題](#)

[方案1.例項ID 1的地理複製校驗和具有IPAM快取和NRFMgmt快取校驗和不匹配](#)

[方案2.例項ID 2的地理複製校驗和與ETCD校驗和不匹配](#)

[案例3.遠端站點的TCP連線建立失敗](#)

[場景4.託管主節點的伺服器上出現DIMM錯誤](#)

[解決方案](#)

簡介

本文檔介紹各種調查方法，用於解決本地機架和遠端機架之間地理複製校驗和不匹配的問題。

必要條件

需求

思科建議您瞭解以下主題：

- 作業階段管理功能(SMF)中的地域備援
- SMF (單一模式光纖)

- 傳輸控制通訊協定(TCP)連線終止

採用元件

本文件所述內容不限於特定軟體和硬體版本。

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除 (預設) 的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

背景資訊

什麼是SMF中的Geo冗餘？

- SMF支援主用 — 主用模式下的地理(Geo)冗餘(GR)。
- GR設定還負責將資料復etcd/cache制到備用機架。
- SMF支援主/備用冗餘，其中資料從主例項複製到備用例項。
- 如果主例項發生故障，備用例項將變為主例項並接管操作。
- 為了實現GR，可以設定兩個主/備用對，其中每個站點都主動處理流量，備用站點充當遠端站點的備份。

異地複製Pod

- 引入了地理複製Pod，用於機架間/站點間通訊以及監控機架內的POD/BFD
- 在每個機架/站點上運行兩個GR-POD例項
- 兩個GR POD在主 — 備用模式下運行
- GR POD在Proto節點/虛擬機器上生成
- GR POD使用兩個虛擬IP地址(VIP)
- 內部VIP，用於POD間通訊 (在機架內)
- 用於機架間/站點的GR POD通訊的外部VIP
- 為GR POD配置的VIP可以在其中一個原始節點/虛擬機器上處於活動狀態
- 當活動GR POD重新啟動時，VIP將切換到另一個Proto節點/VM，而在另一個Proto節點/VM上運行的備用GR POD可以變為活動

GR Pod參考配置：

<#root>

```
smf# show running-config instance instance-id 1 endpoint geo
```

```
Thu Oct 20 06:25:25.319 UTC+00:00
instance instance-id 1
endpoint geo
replicas 1
nodes 2
interface geo-internal
vip-ip a.b.c.d vip-port 7001
exit
interface geo-external
vip-ip Y.Y.Y.Y vip-port 7002
exit
exit
exit
```

確定活動的Geo Pod和備用的Geo Pod

為了識別活動的Geo Pod，您需要檢查Geo Pod日誌中是否存在錯誤或事件。

活動Pod:

<#root>

```
user@smf-ims-master-1:~$ kubectl logs georeplication-pod-0 -n smf-smfix1|tail -3
```

```
[ERROR] [grcacachepod.go:339] [gr_deferred_sync.application.app] Periodic Sync: Total time taken to syn
[ERROR] [GeoAdminStreamClient.go:276] [gr_pod.geo_admin_client.app] no one waiting for received respons
```

備用Pod:

<#root>

```
user@cp0xxx-smf-ims-master-1:~$ kubectl logs georeplication-pod-1 -n smf-smfix1|tail -3
```

```
[ERROR] [gr_pod.geo_replication_client_stream] Counters => not an active geo pod
[ERROR] [gr_pod.geo_replication_client_stream] Counters => not an active geo pod
[ERROR] [gr_pod.geo_replication_client_stream] Counters => not an active geo pod
```

GR POD的功能

GR Pod跨站點複製ETCD和快取Pod資料

要檢視ETCD和快取Pod資料的複製詳細資訊，請使用CLI:

<#root>

```
[cp0xxx-smf-ims/smf1] smf# show georeplication checksum instance-id 1
```

```
Thu Oct 20 07:11:52.409 UTC+00:00
```

```
checksum-details
```

```
--      --
ID      Type      Checksum
--      --
1       ETCD        1666249907
        IPAM CACHE  1666249907
        NRFMgmt CACHE  1666249907
```

在ETCD中維護站點本地例項角色

```
[ERROR] [gr_pod.gradmin] updateEntryInEtcd: Updating etcd entries for keys : Instance.2, with role as P
[ERROR] [gr_pod.gradmin] updateEntryInEtcd: Updating etcd entries for keys : Instance.1, with role as S
```

監控本地站點狀態 (POD狀態/BFD狀態)

```
<#root>
```

```
[cp0xxx-smf-ims/smf1] smf# show running-config geomonitor podmonitor pods smf-service
```

```
Thu Oct 20 07:36:41.280 UTC+00:00
```

```
geomonitor podmonitor pods smf-service
retryCount          2
retryInterval       900
retryFailOverInterval 500
failedReplicaPercent 60
```

站點角色

PRIMARY :站點已準備就緒，可以主動接收給定例項的流量。

STANDBY:站點處於備用狀態，可以接收流量，但是不接收給定例項的流量。

STANDBY_ERROR:站點有問題，處於非活動狀態，且未準備好接收給定例項的流量。

FAILOVER_INIT:站點已啟動故障轉移，並且沒有條件接收流量，緩衝時間為2s，以便應用程式完成其活動。

FAILOVER_COMPLETE:站點已完成故障轉移，並嘗試向對等站點通知給定例項的故障轉移。緩衝時間為2s。

FAILBACK_STARTED:對於給定例項，從遠端站點觸發手動故障切換時會發生延遲。



附註：即使在所有角色中也會發生快取/ETCD複製和CDL複製。如果GR鏈路處於關閉狀態/週期性心跳失敗，則會掛起GR觸發器。

GR-Triggers

用於驗證機架上的GR例項角色的CLI

```
<#root>
```

```
Show role instance id 1
```

```
Show role instance id 2
```

將角色從備用錯誤重置為備用的CLI

```
<#root>
```

```
Geo reset-role instance-id <1/2> role standby
```

CLI to Switch Role from Standby to Standby錯誤

```
<#root>
```

```
Geo switch-role instance-id <1/2> role standby failback-interval 0
```

CLI將角色從備用狀態切換到主狀態

為了啟動此交換機角色，您需要從將其中一個例項作為主例項的機架觸發CLI。

```
<#root>
```

```
Geo switch-role instance-id <1/2> role standby failback-interval 0
```



附註：晴天場景：Rack1-Instance1-Primary、Instance2-Standby;Rack2-Instance1-

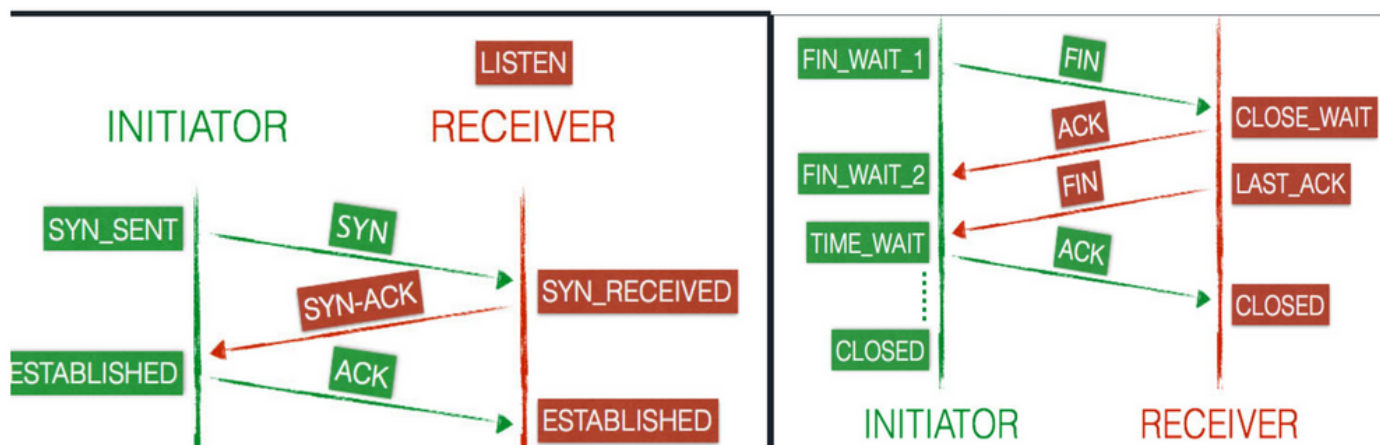
Standby、Instance2-Primary。

Rainy Day場景：機架1例項1和例項2主機架；機架2例項1和例項2備用機架。

TCP連線終止

TCP協定是一種面向連線的協定，這意味著建立和維護連線，直到每一端的應用程式完成消息交換。TCP可搭配網際網路通訊協定(IP)使用。

TCP握手也稱為3次握手。當從客戶端電腦發起到伺服器電腦的連線時，客戶端和伺服器在傳輸資料之前交換SYN和ACK資料包。



傳輸控制協定：客戶端和伺服器連線狀態

連線在其整個生命週期中經歷一系列狀態。這些狀態包括：LISTEN、SYN-SENT、SYN-RECEIVED、ESTABLISHED、FIN-WAIT-1、FIN-WAIT-2、CLOSE-WAIT、CLOSING、LAST-ACK、TIME-WAIT、CLOSED、和虛構狀態。

- SYN開啟新的TCP連線時，客戶端（啟動器）向伺服器（接收器）傳送一個資料包，並將其狀態更新為SYN-SENT TCP。
- 然後，伺服器向客戶端傳送一個回覆，該客戶端將其連線狀態更改為SYN-RECEIVED。
- 客戶端使用ACK作出回覆，且連線在兩個端點上都標籤為ESTABLISHED，此時客戶端和伺服器已準備好傳輸資料。
- 客戶端將資料FIN包傳送到伺服器並將其狀態更新為FIN-WAIT-1。
- 伺服器接收來自客戶端的終止請求並使用ACK響應。回覆後，伺服器進入狀CLOSE-WAIT態。
- 客戶端收到伺服器的回覆後，便會進入狀FIN-WAIT-2態。
- 伺服器仍處於狀CLOSE-WAIT態，它獨立使用FIN，後者將狀態更新為LAST-ACK。
- 現在，客戶端收到終止請求並以ACK回覆，這將導致狀TIME-WAIT態。
- 伺服器現已完成並將連線立即設定CLOSED為。
- 在連線之前，TIME-WAIT客戶端最多會保持四分鐘的狀CLOSED態。

問題

案例1.實例ID 1的地理複製校驗和具有IPAM快取和NRFMgmt快取校驗和不匹配

smfix1/smfix2地域複製狀態失敗 (到遠端站點的機架間複製失敗)。

錯誤：管理命令失敗[pod internal-gr-pod-1,URL <http://X.X.0.0:15290/commands>]，代碼為424，消息失敗：複製校驗和不匹配。

在8月23日00:36:19觀察到此問題，原因是「機架間複製失敗」。

From CEE alerts:

```
Inter_Rack_Replication 9ca45362a049 critical 08-23T00:36:19 System
Inter rack replication to Remote Site failed
```

從此CLI輸出中，您可以看到instance-id 1的Checksum Mismatch for IP Address Management(IPAM)和NRF Cache。

<#root>

```
[cp0xxx-smf-ims/smfix1] smf# show georeplication checksum instance-id 1
```

```
Mon Sep 5 08:38:27.762 UTC+00:00
```

```
checksum-details
```

```
--  ---  -----
ID   Type           Checksum
--  ---  -----
1    ETCD           1662367102
      IPAM  CACHE  1662367102
      NRFMgmtCACHE 1662367102
```

<#root>

```
[cp0xxx-smf-ims/smfix2] smf# show georeplication checksum instance-id 1
```

```
Mon Sep 5 08:38:30.767 UTC+00:00
```

```
checksum-details
```

```
--  ---  -----
ID   Type           Checksum
--  ---  -----
1    ETCD           1662367102
      IPAM  CACHE  1661214831
      NRFMgmtCACHE 1661214831
```

案例2.實例ID 2的地理複製校驗和與ETCD校驗和不匹配

```
[cp0xxx-smf-ims/smf1] smf# show georeplication checksum instance-id 2
Mon Sep 5 08:38:37.852 UTC+00:00
checksum-details
--      ----      -----
ID      Type          Checksum
--      ----      -----
2       ETCD          1661214828
        IPAM    CACHE  1662367107
        NRFMgmtCACHE 1662367107
```

<#root>

```
[cp0xxx-smf-ims/smf2] smf# show georeplication checksum instance-id 2

Mon Sep 5 08:38:39.118 UTC+00:00
checksum-details
--      ----      -----
ID      Type          Checksum
--      ----      -----
2       ETCD          1662367107
        IPAM    CACHE  1662367107
        NRFMgmtCACHE 1662367107
```

案例3.與遠程站點建立TCP連線失敗

Rack1-smfix1-logs:

從GR Pod日誌，您可以觀察到更新快取Pod檢查點已停止、立即複製失敗以及沒有可用的遠端主機

。

```
2022/08/23 00:34:00.035 [ERROR] [grreplicationclient.go:201] [gr_pod.geo_replication_client_stream.app]
2022/08/23 00:34:02.086 [ERROR] [grreplicationclient.go:466] [gr_pod.geo_replication_client_stream.app]
2022/08/23 00:34:04.124 [ERROR] [GeoAdminStreamClient.go:215] [gr_pod.geo_admin_client.app] ADMIN(geo-
2022/08/23 00:34:43.623 [ERROR] [grreplicationclient.go:270] [gr_pod.geo_replication_client_stream.app]
```

Rack2-smfix2-logs:

從GR Pod日誌，您可以觀察到流斷開連線錯誤和CACHE校驗和差異超出預期。

```
2022/08/23 00:34:06.497 [ERROR] [grreplicationserver.go:62] [gr_pod.geo_replication_server_stream.app]
2022/08/23 00:34:06.497 [ERROR] [grreplicationserver.go:314] [gr_pod.geo_replication_server_stream.app]
2022/08/23 00:34:56.751 [ERROR] [grpodcommands.go:455] [gr_pod.cli_command.app] compareChecksumData: C
2022/08/23 00:34:56.678 [ERROR] [etcdAuditReplHandler.go:196] [gr_pod.application.app] SyncETCDData pe
2022/08/23 00:36:56.757 [ERROR] [grpodcommands.go:455] [gr_pod.cli_command.app] compareChecksumData: C
```

場景4.託管主節點的伺服器上出現DIMM錯誤

大約在流斷開連線錯誤的同時，主節點1上出現ECC錯誤，該節點承載geo-replication-pod-0。

<#root>

```
CP0XXX-Server9-02# scope sel
```

```
CP0XXX-Server9-02 /sel # show entries
```

Time	Severity	Description
2022-08-23 00:33:59 UTC	Informational	"DDR4_P1_E1_ECC: Memory sensor, read 1 correctable ECC errors on
2022-08-22 22:59:45 UTC	Informational	"DDR4_P1_E1_ECC: Memory sensor, read 1 correctable ECC errors on

- 機架1上的Geo-replication-pod與機架2上的Geo-replication-pod之間的通訊中斷。
- 一個主節點上出現DIMM錯誤，導致機架1和機架2之間的資料流連線中斷。
- 從Rack1 Geo-replication-pod無法複製任何請求或向Rack2傳送任何請求，出現錯誤Remote Host not available。
- 從Rack1和Rack2上針對7002埠的netstat命令輸出中發現，Rack1插槽停滯在FIN_WAIT1狀態，而Rack2插槽停滯在SYN_RECV狀態。
- 在伺服器端，即Rack2上，套接字停滯在SYNC_RECV狀態，並且新建立的連線也處於SYNC_RECV狀態，並且無法相互通訊。
- 連線處於SYN_RECV狀態，因為核心已收到某個埠的SYN資料包（即處於偵聽模式），但另一端沒有使用ACK進行應答。

smfix2-Master-2已安裝geo外部VIP(Y.Y.Y.Y:7002)，但遠端主機(SMFX1)TCP連線狀態停滯在SYN_RECV狀態而不是ESTABLISHED狀態。a.b.c.d和a.b.c.e是smfix1（機架1）的Master-1和2 IP。

<#root>

```
user@cp0xxx-smf-ims-master-2:~$ netstat -anp | grep 7002
```

tcp	0	0	Y.Y.Y.Y:7002	0.0.0.0:*	LISTEN	-
tcp	0	0	Y.Y.Y.Y:7002	a.b.c.e:35542	SYN_RECV	-
tcp	0	0	Y.Y.Y.Y:7002	a.b.c.d:47046	SYN_RECV	-
tcp	0	0	Y.Y.Y.Y:7002	a.b.c.e:36248	SYN_RECV	-
tcp	0	0	Y.Y.Y.Y:7002	a.b.c.d:42686	SYN_RECV	-
tcp	0	0	Y.Y.Y.Y:7002	a.b.c.e:38248	SYN_RECV	-

遠端對等體的smfix1 (機架1) 上的外部Geo VIP TCP連線狀態處於FIN-WAIT1狀態 :

```
<#root>
```

```
user@cp0xxx-smf-ims-master-1:~$ netstat -anp | grep 7002
```

tcp	0	0	a.b.c.d	0.0.0.0:*	LISTEN	-
tcp	0	1	a.b.c.d:60866	Y.Y.Y.Y:7002	FIN_WAIT1	-
tcp	0	1	a.b.c.d:52274	Y.Y.Y.Y:7002	FIN_WAIT1	-
tcp	0	1	a.b.c.d:59674	Y.Y.Y.Y:7002	FIN_WAIT1	-
tcp	0	1	a.b.c.d:47926	Y.Y.Y.Y:7002	FIN_WAIT1	-

解決方案

機架1:

- 首先，刪除備用Geo Pod，等待Pod恢復，然後刪除活動Geo Pod。登入到主VIP並刪除GR Pod:

```
<#root>
```

```
kubectl delete pod
```

-n

機架2:

- 首先，刪除備用Geo Pod，等待Pod恢復，然後刪除Active Geo Pod。
- 從CLI檢查地理複製狀態，然後刪除地理位置pod。

```
<#root>
```

```
show georeplication-status
```

- 在Rack1和Rack2上刪除Geo Pod，您可以看到外部Geo VIP IP:TCP埠將變為

ESTABLISHED狀態。

- GeoReplication狀態「通過」。
- 在機架間的複製狀態中未發現校驗和不匹配。

smfix2 (機架2) :

<#root>

```
user@cp0xxx-smf-ims-master-1:~$ sudo netstat -anp | grep 7002 | grep -v aa
```

```
tcp      0      0 Y.Y.Y.Y:7002      0.0.0.0:*          LISTEN      36854
tcp      0      0 Y.Y.Y.Y:7002      a.b.c.d:46402      ESTABLISHED 36854/grpod
tcp      0      0 Y.Y.Y.Y:7002      1a.b.c.e:54708     ESTABLISHED 36854/grpod
tcp      0      0 Y.Y.Y.Y:7002      a.b.c.d:55152     ESTABLISHED 36854/grpod
tcp      0      0 Y.Y.Y.Y:7002      a.b.c.e:46530     ESTABLISHED 36854/grpod
tcp      0      0 10.59.0.0:7002    10.59.0.0:46532   ESTABLISHED 36854/grpod
```

smfix1 (機架1) :

```
user@cp0xxx-smf-ims-master-1:~$ sudo netstat -anp | grep 7002 | grep -v aa
```

```
tcp      0      0 a.b.c.d           0.0.0.0:*          LISTEN      53932/grpod
tcp      0      0 a.b.c.d:46530     Y.Y.Y.Y:7002      ESTABLISHED 53932/grpod
tcp      0      0 a.b.c.d:46402     Y.Y.Y.Y:7002      ESTABLISHED 53932/grpod
tcp      0      17 a.b.c.d:46532     Y.Y.Y.Y:7002      ESTABLISHED 53932/grpod
```

2.異地複製狀態 :

<#root>

```
[okcp0xx-smf-ims/smfix1] smf# show georeplication-status
```

```
result "pass"
```

```
[okcp0xx-smf-ims/smfix2] smf# show georeplication-status
```

```
result "pass"
```

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。