

# 使用错误代码"424-Geo-replication Checksum Mismatch"；排除机架间复制故障

## 目录

---

### [简介](#)

### [先决条件](#)

#### [要求](#)

#### [使用的组件](#)

### [背景信息](#)

#### [什么是SMF中的Geo冗余？](#)

#### [地区复制Pod](#)

#### [确定活动的Geo Pod和备用的Geo Pod](#)

#### [GR POD的功能](#)

##### [GR Pod跨站点复制ETCD和缓存Pod数据](#)

##### [在ETCD中维护站点本地实例角色](#)

##### [监控本地站点状态（POD状态/BFD状态）](#)

#### [站点角色](#)

#### [GR触发器](#)

##### [用于验证机架上的GR实例角色的CLI](#)

##### [CLI to Reset Role from Standby Error to Standby](#)

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

##### [CLI将角色从备用设备切换为主设备](#)

#### [TCP连接终止](#)

### [问题](#)

#### [方案1.实例ID 1的地理复制校验和IPAM缓存和NRFMgmt缓存校验和不匹配](#)

#### [方案2.实例ID 2的地理复制校验和与ETCD校验和不匹配](#)

#### [场景3.与远程站点建立TCP连接失败](#)

#### [场景4.在托管主节点的服务器上观察到的DIMM错误](#)

### [解决方案](#)

---

## 简介

本文档介绍各种调查方法，用于解决本地和远程机架间的地理复制校验和不匹配问题。

## 先决条件

### 要求

Cisco 建议您了解以下主题：

- 会话管理功能(SMF)中的地域冗余
- SMF

- 传输控制协议(TCP)连接终止

## 使用的组件

本文档不限于特定的软件和硬件版本。

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始 ( 默认 ) 配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

## 背景信息

### 什么是SMF中的Geo冗余？

- SMF支持主用 — 主用模式中的地理(Geo)冗余(GR)。
- GR设置还负责将数据复etcd/cache制到备用机架。
- SMF支持主/备用冗余，其中数据从主实例复制到备用实例。
- 如果主实例发生故障，备用实例将变为主实例并接管操作。
- 要实现GR，可以设置两个主/备用对，其中每个站点主动处理流量，备用作为远程站点的备份。

### 地区复制Pod

- 引入了地理复制Pod，用于机架间/站点间通信以及监控机架内的POD/BFD
- 每个机架/站点上运行两个GR-POD实例
- 两个GR POD在主用 — 备用模式下运行
- GR POD在原始节点/虚拟机上生成
- GR POD使用两个虚拟IP地址(VIP)
- 内部VIP，用于POD间通信 ( 在机架内 )
- 用于机架间/站点间GR POD通信的外部VIP
- 为GR POD配置的VIP可以在其中一个原始节点/虚拟机上处于活动状态
- 当主用GR POD重新启动时，VIP会切换到另一个Proto节点/VM，而在其他Proto节点/VM上运行的备用GR POD可以变为主用

GR Pod参考配置：

<#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
```

## 确定活动的Geo Pod和备用的Geo Pod

为了识别活动的Geo Pod，您需要检查Geo Pod日志中的错误或事件。

活动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
```

备用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跨站点复制ETCD和缓存Pod数据

要查看ETCD和缓存Pod数据的复制详细信息，请使用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
        NRMgmt CACHE 1666249907
```

## 在ETCD中维护站点本地实例角色

```
[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
```

## 监控本地站点状态 ( POD状态/BFD状态 )

```
<#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
```

## 站点角色

**PRIMARY** :站点已准备好并主动接收给定实例的流量。

**STANDBY**:站点处于备用状态，可以接收流量，但不会接收给定实例的流量。

**STANDBY\_ERROR**:站点有问题、未激活以及未准备好接收给定实例的流量。

**FAILOVER\_INIT**:站点已开始进行故障切换，并且不处于接收流量的状态，缓冲时间为2s，应用才能完成其活动。

**FAILOVER\_COMPLETE**:站点已完成故障转移，并尝试向对等站点通知给定实例的故障转移。2秒的缓冲时间。

**FAILBACK\_STARTED**:手动故障切换从远程站点触发并延迟给定实例。



注意：即使在所有角色中也会发生缓存/ETCD复制和CDL复制。如果GR链路发生故障/定期心跳故障，则会挂起GR触发器。

---

## GR触发器

用于验证机架上的GR实例角色的CLI

```
<#root>
```

```
Show role instance id 1
```

```
Show role instance id 2
```

CLI to Reset Role from Standby Error to Standby

```
<#root>
```

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

CLI to Switch Role from Standby to Standby错误

```
<#root>
```

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

CLI将角色从备用设备切换为主设备

要启动此交换机角色，您需要从将其中一个实例作为主实例的机架触发CLI。

```
<#root>
```

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



注意：晴天场景：Rack1-Instance1-Primary、Instance2-Standby;Rack2-Instance1-

---

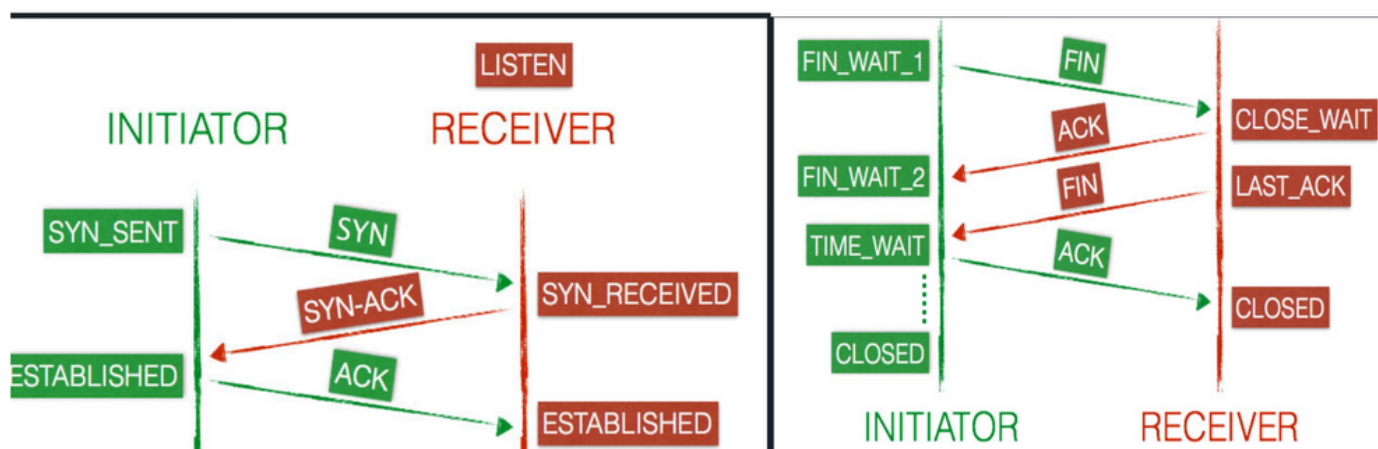
StandBy、Instance2-Primary。

雨天情景：Rack1-Instance 1和Instance 2-Primary;Rack2-Instance 1和Instance 2-StandBy。

## TCP连接终止

TCP协议是一种面向连接的协议，这意味着建立并维护连接，直到每一端的应用程序完成消息交换。TCP与Internet协议(IP)配合使用。

TCP握手也称为三次握手。当从客户端计算机发起到服务器计算机的连接时，客户端和服务器在传输数据之前交换SYN和ACK数据包。



传输控制协议：客户端和服务器连接状态

连接在其整个生命周期中经历一系列状态。状态如下：LISTEN、SYN-SENT、SYN-RECEIVED、ESTABLISHED、FIN-WAIT-1、FIN-WAIT-2、CLOSE-WAIT、CLOSING、LAST-ACK、TIME-WAIT、CLOSED、和虚拟状态。

- SYN当新的TCP连接打开时，客户端（发起方）向服务器（接收方）发送数据包，并将其状态更新为SYN-SENT TCP。
- 然后，服务器向客户端发送一个应答，将其连接状态更改为SYN-RECEIVED。
- 客户端回复ACK且连接在两个终端上都标记为ESTABLISHED，此时客户端和服务器已准备好传输数据。
- 客户端将数据FIN包发送到服务器并将其状态更新为FIN-WAIT-1。
- 服务器接收来自客户端的终止请求并以ACK响应。应答后，服务器进入CLOSE-WAIT态。
- 客户端收到服务器的回复后，便会进入FIN-WAIT-2态。
- 服务器仍处于CLOSE-WAIT态，它独立与FIN通信，FIN将状态更新为LAST-ACK。
- 现在，客户端收到终止请求并回复ACK一个，从而产生TIME-WAIT态。
- 服务器现已完成并将连接立即设置CLOSED为。
- 客户端在连接之前最多保持四分钟的CLOSED态。

## 问题

## 方案1.实例ID 1的地理复制校验和存在IPAM缓存和NRFMgmt缓存校验和不匹配问题

smfix1/smfix2异地复制状态为失败（到远程站点的机架间复制失败）。

错误：管理命令失败[pod internal-gr-pod-1, URL <http://X.X.0.0:15290/commands>]，代码为424，消息失败：复制校验和不匹配。

在8月23日00:36:19观察到此问题为“机架间复制失败”。

From CEE alerts:

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

从此CLI输出中，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
```

## 方案2.实例ID 2的地理复制校验和与ETCD校验和不匹配

```
[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
```

### 场景3.与远程站点建立TCP连接失败

Rack1-smfix1-logs:

从GR Pod日志，您可以观察到更新缓存Pod检查点已停止、立即复制失败以及没有可用的远程主机。

```
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:

从GR Pod日志中，您可以观察到数据流断开连接错误，以及CACHE校验和差异超出预期。

```
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
```

#### 场景4.在托管主节点的服务器上观察到的DIMM错误

在大约与数据流断开错误同时托管geo-replication-pod-0的master-1节点上出现ECC错误。

<#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 (

- Rack1上的Geo-replication-pod和Rack2上的Geo-replication-pod之间的通信中断。
- DIMM错误发生在其中一个主节点上，导致机架1和机架2之间的数据流连接断开。
- 从Rack1 Geo-replication-pod无法复制任何请求或向Rack2发送任何请求，结果显示错误 Remote Host not available。
- 从Rack1和Rack2上针对7002端口的netstat命令输出中发现，Rack1插槽停滞在FIN\_WAIT1状态，而Rack2插槽停滞在SYN\_RECV状态。
- 在服务器端，即Rack2上，套接字停滞在SYNC\_RECV状态，新创建的连接也进入 SYNC\_RECV状态，且无法相互通信。
- 连接处于SYN\_RECV状态，因为内核已收到一个端口的SYN数据包（即，在LISTENING模式下），但另一端没有使用ACK进行应答。

smfix2-Master-2已安装geo外部VIP(Y.Y.Y.Y:7002)，但远程主机(SMFIX1)TCP连接状态停滞在SYN\_RECV状态而不是ESTABLISHED状态。a.b.c.d和a.b.c.e是smfix1（机架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	-

远程对等体smfix1 ( 机架1 ) 上的外部Geo VIP TCP连接状态处于FIN-WAIT1状态 :

<#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	-

## 解决方案

机架1:

- 首先，删除备用Geo Pod，等待Pod恢复，然后删除活动Geo Pod。登录主VIP并删除GR Pod:

<#root>

```
kubectl delete pod
```

-n

机架2:

- 首先，删除备用Geo Pod，等待Pod恢复，然后删除Active Geo Pod。
- 从CLI检查地理复制状态，然后删除地理位置Pod。

<#root>

```
show georeplication-status
```

- 在Rack1和Rack2上删除Geo Pod，您可以看到外部Geo VIP IP:TCP端口将变为

ESTABLISHED状态。

- GeoReplication状态“通过”。
- 在机架间的复制状态中看不到校验和不匹配。

smfix2 ( 机架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 ( 机架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.地理复制状态 :

<#root>

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

```
result "pass"
```

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

```
result "pass"
```

## 关于此翻译

思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。