Solución de problemas de transmisión de la comunidad de salida RMAC EVPN al fabric ACI

Contenido

Introducción

Este documento describe el impacto del atributo de comunidad ampliada MAC del router mal configurado en un fabric ACI cuando se recibe de un peer BGP (Border Gateway Protocol) externo.

Antecedentes

Con BGP, existe la opción de enviar atributos de comunidad ampliada y de comunidad con los prefijos que se anuncian a los peers BGP. Estos atributos de comunidad nos permiten modificar las políticas de ruteo y alterar dinámicamente la manera en que se maneja el tráfico ruteado.

Problema

Cuando el atributo de comunidad ampliada MAC del router se envía con un prefijo AFI IPv4 desde un peer BGP externo a un fabric ACI, se produce un error de programación FIB y HAL en cualquier hoja del fabric que reciba la ruta desde las hojas de borde a través del proceso MP-BGP interno. Esto se debe a que el atributo extcommunity de RMAC pertenece a la familia de direcciones BGP L2VPN EVPN y, cuando se inyecta en la familia de direcciones BGP IPv4, se rechaza. Esto se debe a una violación de la regla 5.2 (Modo de Propagación Uniforme), que se describe en el documento de IETF titulado "Interacción EVPN con IPVPN". En la página 15, en el punto 4c, se señala la cuestión específica:

```
4.  As discussed, Communities, Extended Communities and Large
      Communities SHOULD be kept by the gateway PE from the originating
      SAFI route.  Exceptions of Extended Communities that SHOULD NOT
      be kept are:

      C.  All the extended communities of type EVPN.

      The gateway PE SHOULD NOT copy the above extended communities
      from the originating ISF route to the re-advertised ISF route.
```

Enlace al documento: [Interacción de EVPN con IPVPN](#)

Aquí hay un ejemplo del problema con iBGP, sin embargo, el problema también se ve con eBGP.

Diagrama de topología:

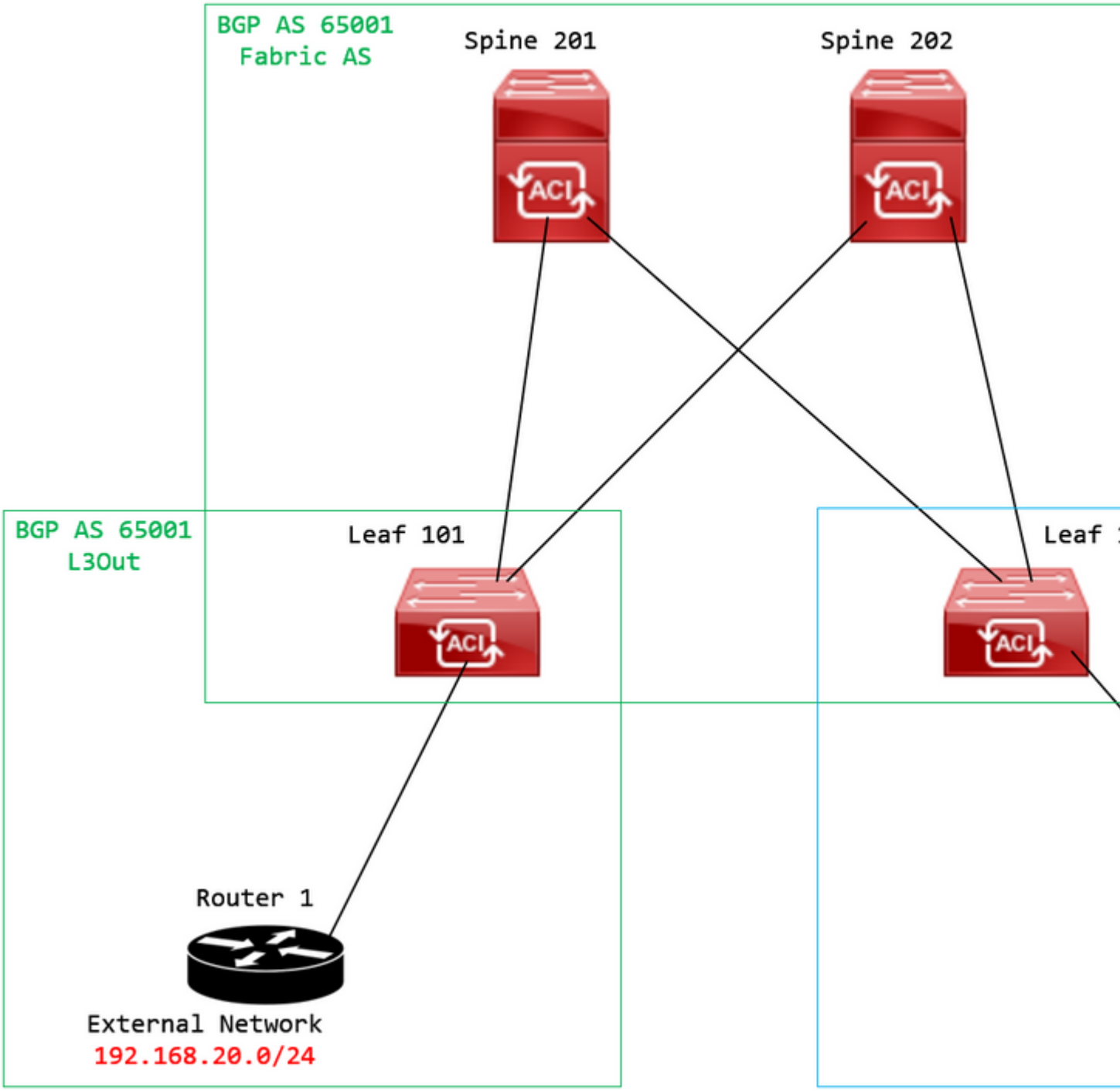


Diagrama de topología

Configure el route map en el dispositivo par BGP externo (Router 1) y establezca el atributo extcommunity de EVPN RMAC:

```
Router-1# show run | sec route-map
route-map RMAC permit 10
  set extcommunity evpn rmac aaaa.bbbb.cccc
```

En la configuración de la familia de direcciones IPv4 del vecino BGP, configure las comunidades extendidas BGP y configure el route map en la dirección saliente:

```
Router-1# show run bgp
```

```
<output omitted>
feature bgp

router bgp 65001
  vrf example
    router-id 192.168.20.20
    address-family ipv4 unicast
      network 192.168.20.0/24
    neighbor 192.168.30.30
      remote-as 65001
      update-source loopback1
```

address-family ipv4 unicast

send-community extended

route-map RMAC out

Verifique el estado de BGP en BL 101:

<#root>

```
leaf-101# show ip bgp 192.168.20.0 vrf example:example
BGP routing table information for VRF example:example, address family IPv4 Unicast
BGP routing table entry for 192.168.20.0/24, version 40 dest ptr 0xa0fec840
Paths: (1 available, best #1)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best urib route, is in HW, exported
  vpn: version 2725, (0x100002) on xmit-list
Multipath: eBGP iBGP

  Advertised path-id 1, VPN AF advertised path-id 1
  Path type (0xa96485b8): internal 0x18 0x0 ref 0 adv path ref 2, path is valid, is best path
  AS-Path: NONE, path sourced internal to AS
    192.168.20.20 (metric 5) from 192.168.20.20 (192.168.20.20)
      Origin IGP, MED not set, localpref 100, weight 0 tag 0, propagate 0
      Extcommunity:
          RT:65001:2162688
          COST:pre-bestpath:163:1879048192
```

Router MAC:aaaa.bbbb.cccc

******Notice that the router mac is present here.******

```
          VNID:2162688

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
```

```
  Path-id 1 advertised to peers:
    10.0.216.65        10.0.216.66
```

Compruebe el NERVIO en CL 102:

<#root>

```
leaf-102# show ip route 192.168.20.0 vrf example:example
IP Route Table for VRF "example:example"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>
```

192.168.20.0/24

```
, ubest/mbest: 1/0
    *via
```

10.0.210.70

```
%overlay-1, [200/0], 00:00:43, bgp-65001, internal, tag 65001,
```

rwVnid: vxlan-2162688

```
         recursive next hop: 10.0.210.70/32%overlay-1
```

******Notice that we have the route here and our next-hop address is correct (showing the TEP IP of BL 101)***

```
leaf-102# acidiag fnvread | grep 101
     101        1             leaf-101      <output omitted>
```

10.0.210.70/32

```
    leaf         active   0
```

Comprobar FIB en CL 102:

<#root>

```
module-1(DBG-elam-insel6)# show forwarding route 192.168.20.0 vrf example:example
ERROR: no longest match in IPv4 table 0xf5df36b0
```

******No entry is present.******

Compruebe la tabla HAL en CL 102:

<#root>

```
module-1(DBG-elam-insel6)# show platform internal hal l3 routes | grep 192.168.20.0
```

******No entry is present.******

Pings de EP (Host 1) al host en una red externa que proviene de un peer BGP externo (192.168.20.20):

<#root>

```
Host-1# ping 192.168.20.20 vrf example
PING 192.168.20.20 (192.168.20.20): 56 data bytes
Request 0 timed out
Request 1 timed out
Request 2 timed out
Request 3 timed out
Request 4 timed out

--- 192.168.20.20 ping statistics ---
5 packets transmitted, 0 packets received, 100.00% packet loss
```

******No connectivity.******

Compruebe ELAM en CL 102:

<#root>

```
leaf-102# vsh_lc
module-1# debug platform internal roc elam asic 0
module-1(DBG-elam)# trigger reset
module-1(DBG-elam)# trigger init in-select 6 out-select 0
module-1(DBG-elam-insel6)# set outer ipv4 src_ip 192.168.10.10 dst_ip 192.168.20.20
module-1(DBG-elam-insel6)# start
module-1(DBG-elam-insel6)# stat
 ELAM STATUS
===========
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Triggered

module-1(DBG-elam-insel6)# ereport
Python available. Continue ELAM decode with LC Pkg
 ELAM REPORT
<output omitted>
--------------------------------------------------------------------------------------------------------------------------------
Lookup Drop
--------------------------------------------------------------------------------------------------------------------------------
LU drop reason                          :
```

UC_PC_CFG_TABLE_DROP

******Notice the drop vector here.******

Solución

La solución consiste en dejar de enviar el atributo de comunidad ampliada MAC del router con un prefijo de familia de direcciones IPv4 desde un par BGP externo a un fabric ACI.

Quite el route map configurado anteriormente y deje de enviar comunidades extendidas desde el dispositivo de peer BGP externo (Router 1). La eliminación de una de estas configuraciones, o de ambas, funcionará:

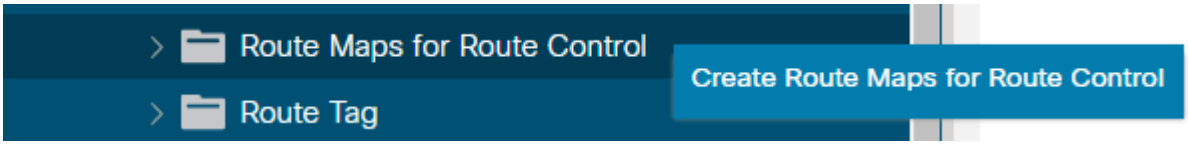
```
Router-1# show run bgp

feature bgp

router bgp 65001
  vrf example
    router-id 192.168.20.20
    address-family ipv4 unicast
      network 192.168.20.0/24
    neighbor 192.168.30.30
      remote-as 65001
      update-source loopback1
      address-family ipv4 unicast
```

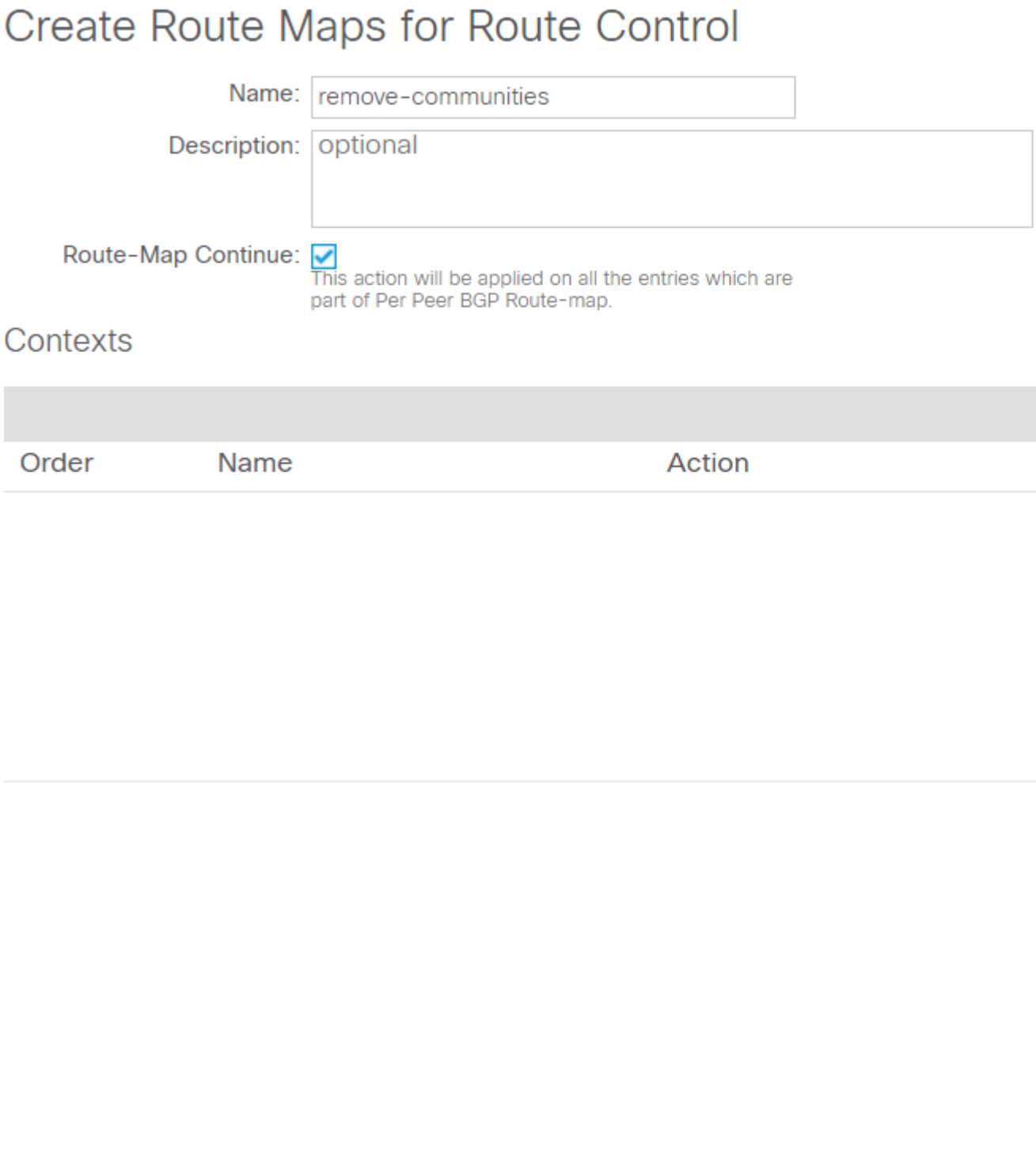
Otra solución (menos preferida) consiste simplemente en filtrar todas las comunidades recibidas del dispositivo de peer BGP externo creando un route map en el L3Out configurado en ACI.

Desplácese hasta su Tenant > Policies > Protocol > Route Maps for Route Control > Create Route Maps for Route Control:



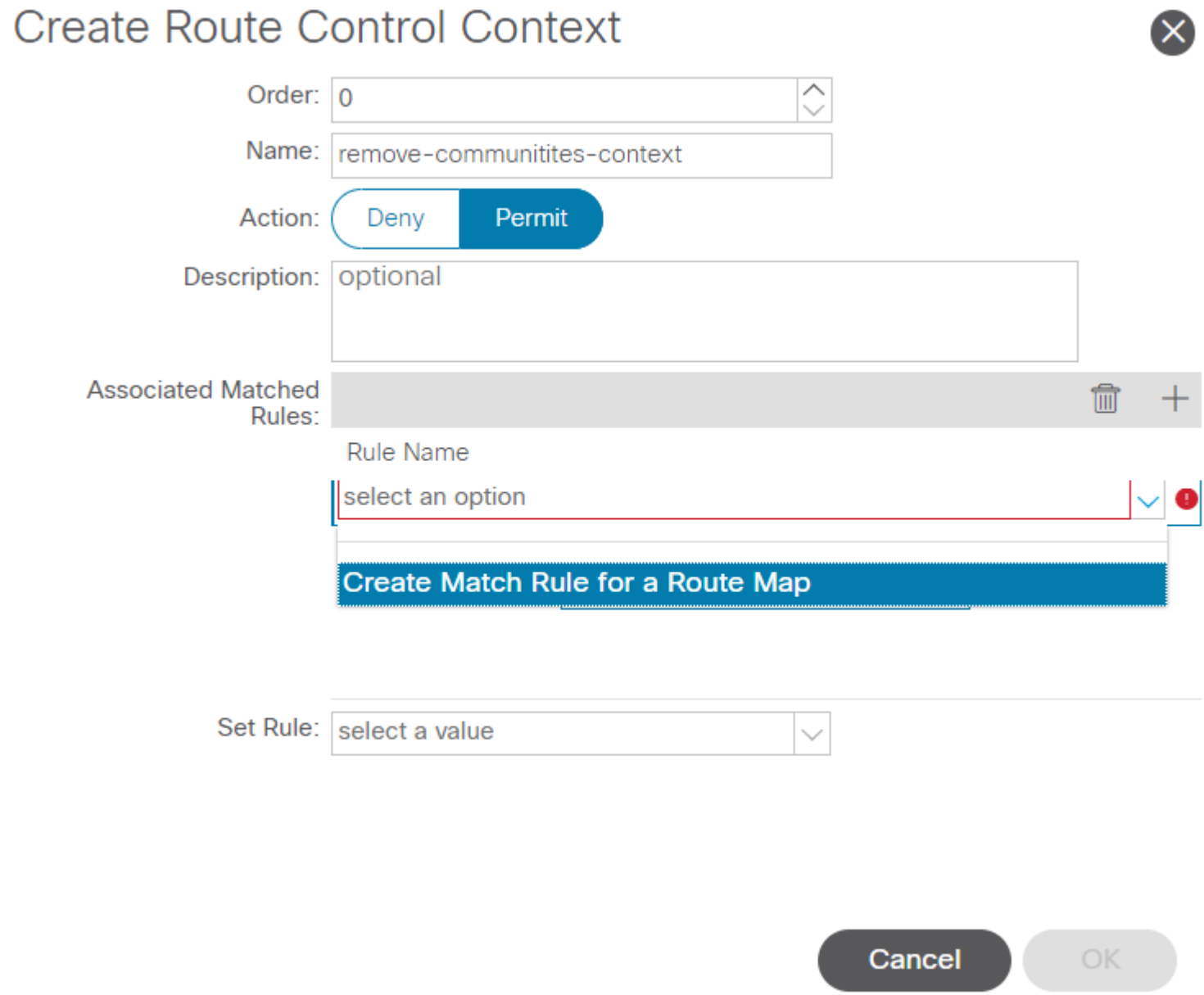
Seleccione la opción Crear mapas de ruta para el control de ruta

Asigne un nombre al mapa de ruta y active la Route-Map Continue y, a continuación, agregue un contexto. Seleccione el + en la tabla **Contextos**:



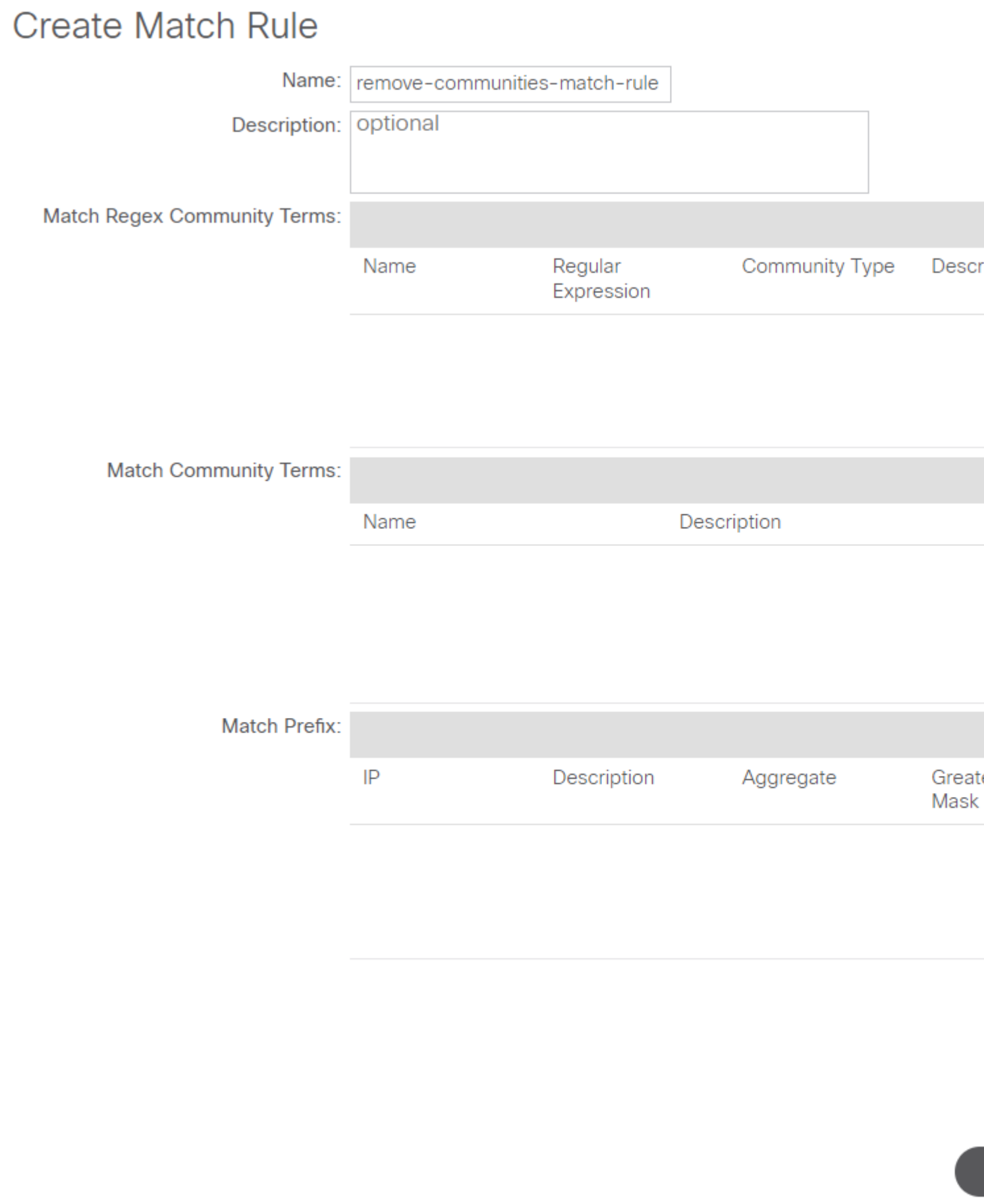
Crear mapa de ruta y crear contexto

Nombre el contexto y deje la acción predeterminada de Permit seleccionado y, a continuación, cree una regla de coincidencia seleccionando el + icono en el Associated Matched Rules tabla y seleccione **Create Match Rule for a Route Map**:



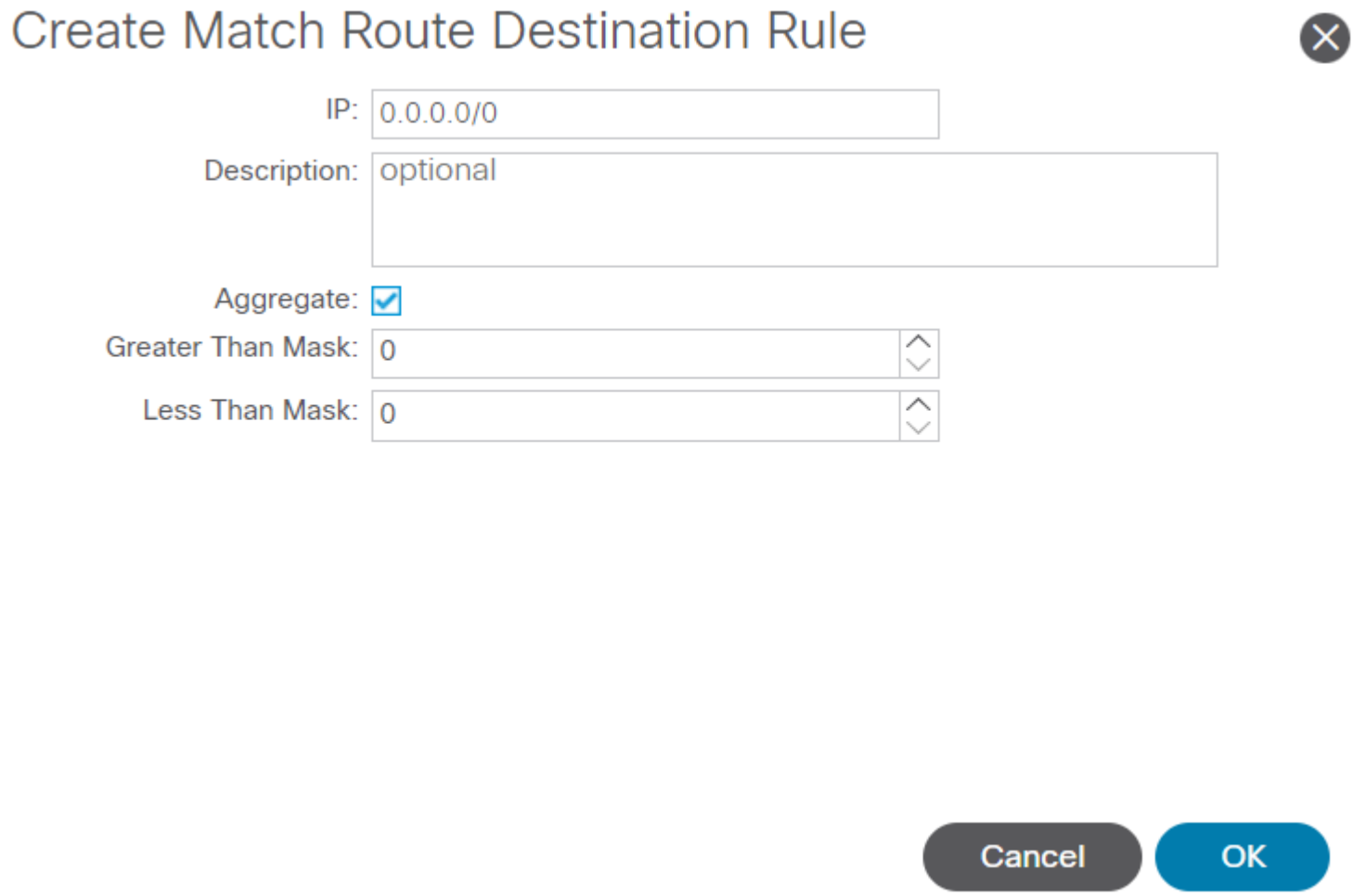
Crear contexto de control de ruta y seleccionar la opción Crear regla de coincidencia para un mapa de ruta

Asigne un nombre a la regla de coincidencia y, a continuación, agregue un nuevo prefijo seleccionando el icono + del Match Prefix tabla:



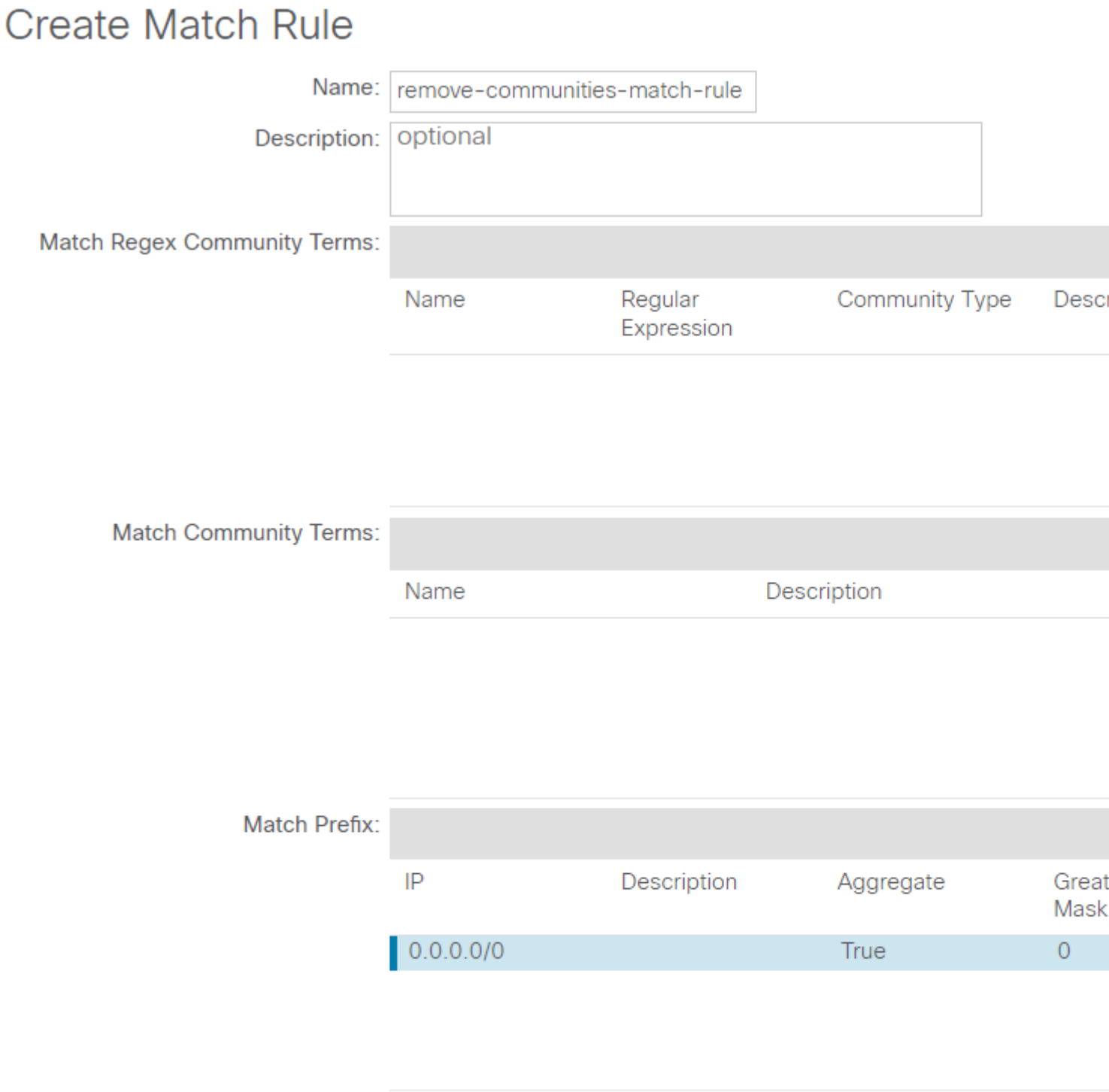
Crear regla de coincidencia y crear prefijo de coincidencia

Agregue el prefijo que desee. Este ejemplo muestra cómo agregar un agregado de todos los prefijos:



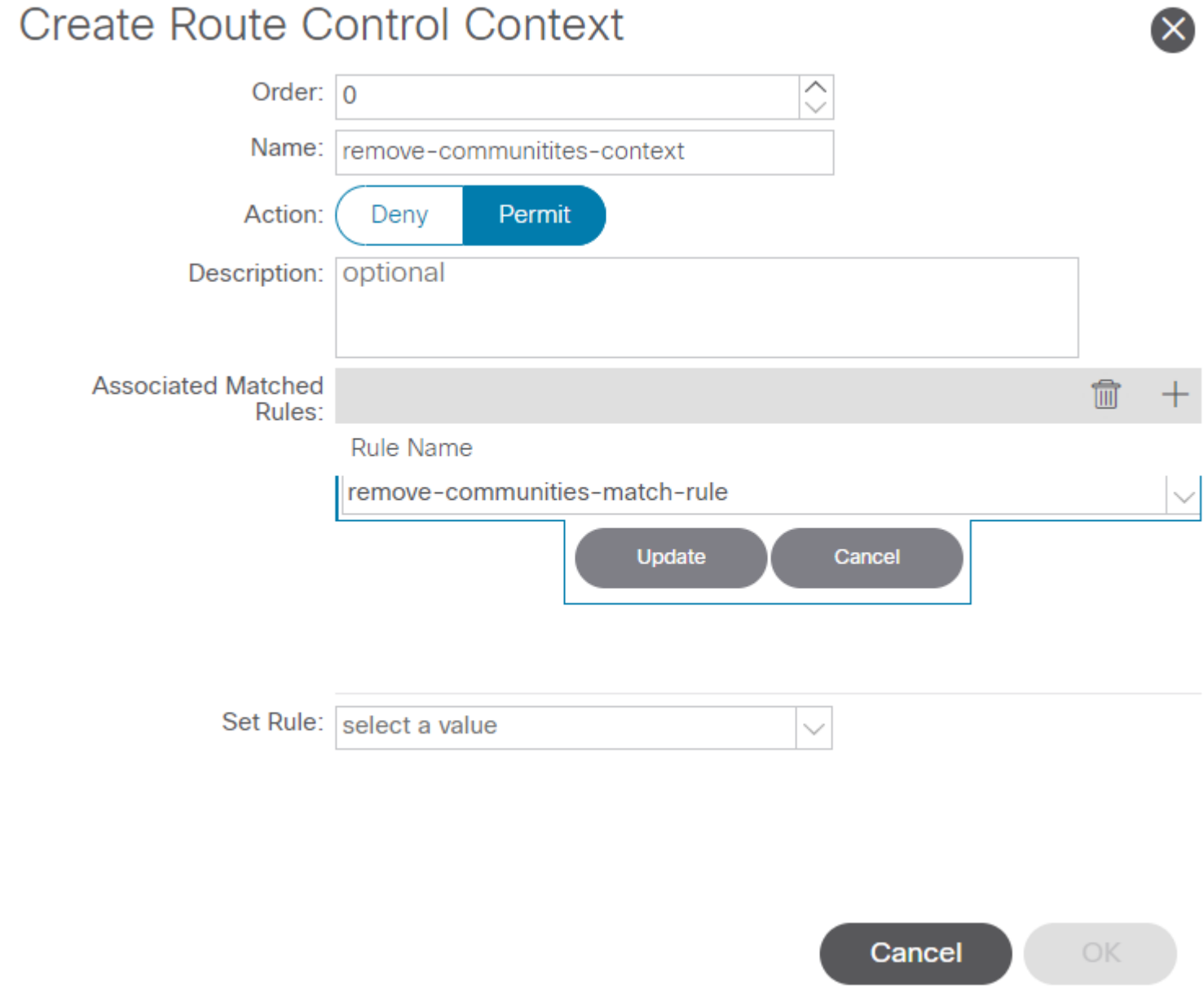
Crear Regla de Destino de Ruta de Coincidencia

Después de seleccionar **OK** en el Create Match Route Destination Rule , verá que el prefijo se ha agregado al Match Prefix tabla en el Create Match Rule ventana:



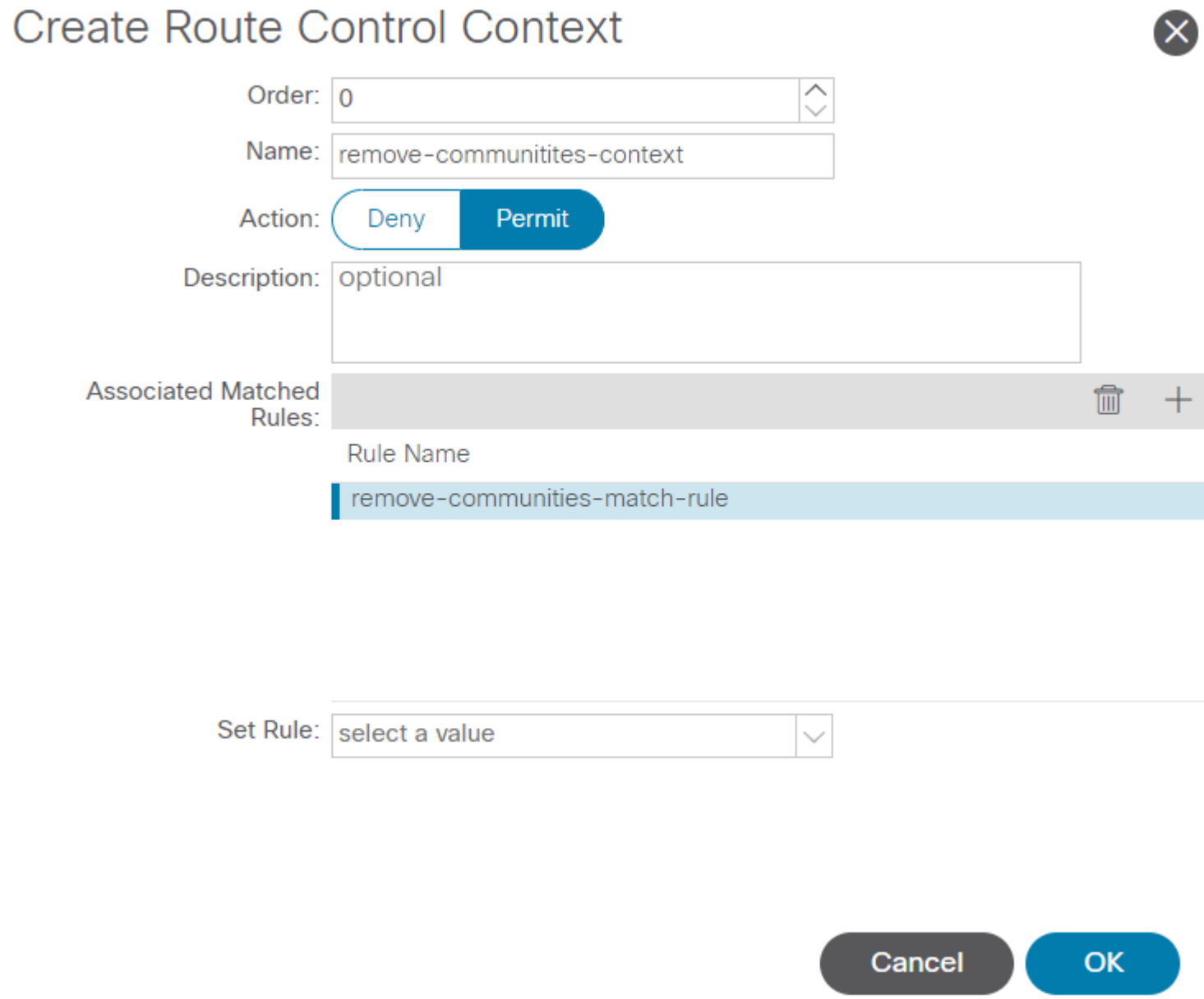
El prefijo de coincidencia se agrega ahora a la regla de coincidencia

Después de seleccionar Submit en el Create Match Rule ventana, seleccione Update en el Associated Matched Rules tabla en el Create Route Control Context ventana:



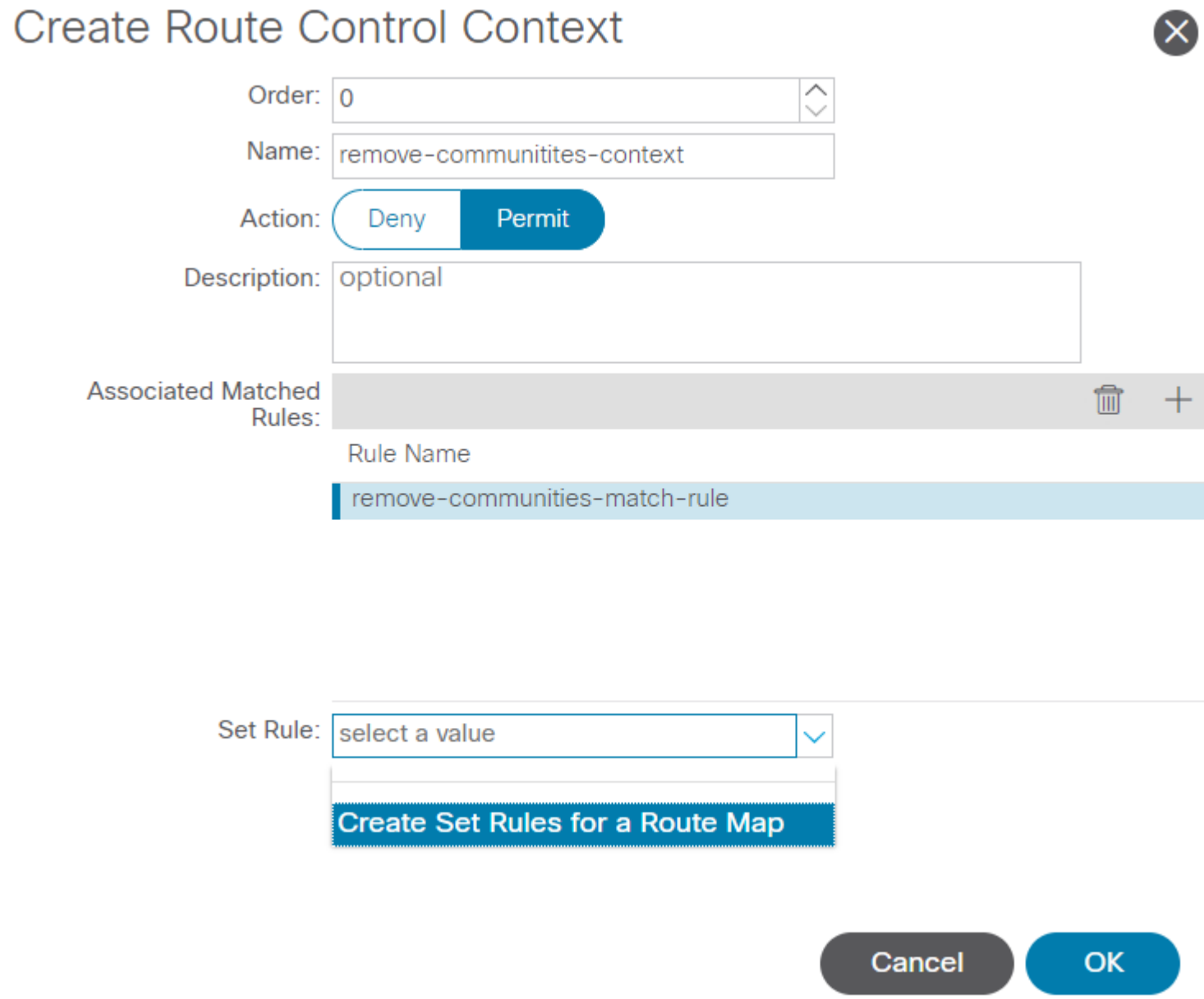
Agregar regla de coincidencia asociada al contexto de control de ruta

La regla de coincidencia asociada se agrega ahora a su contexto:



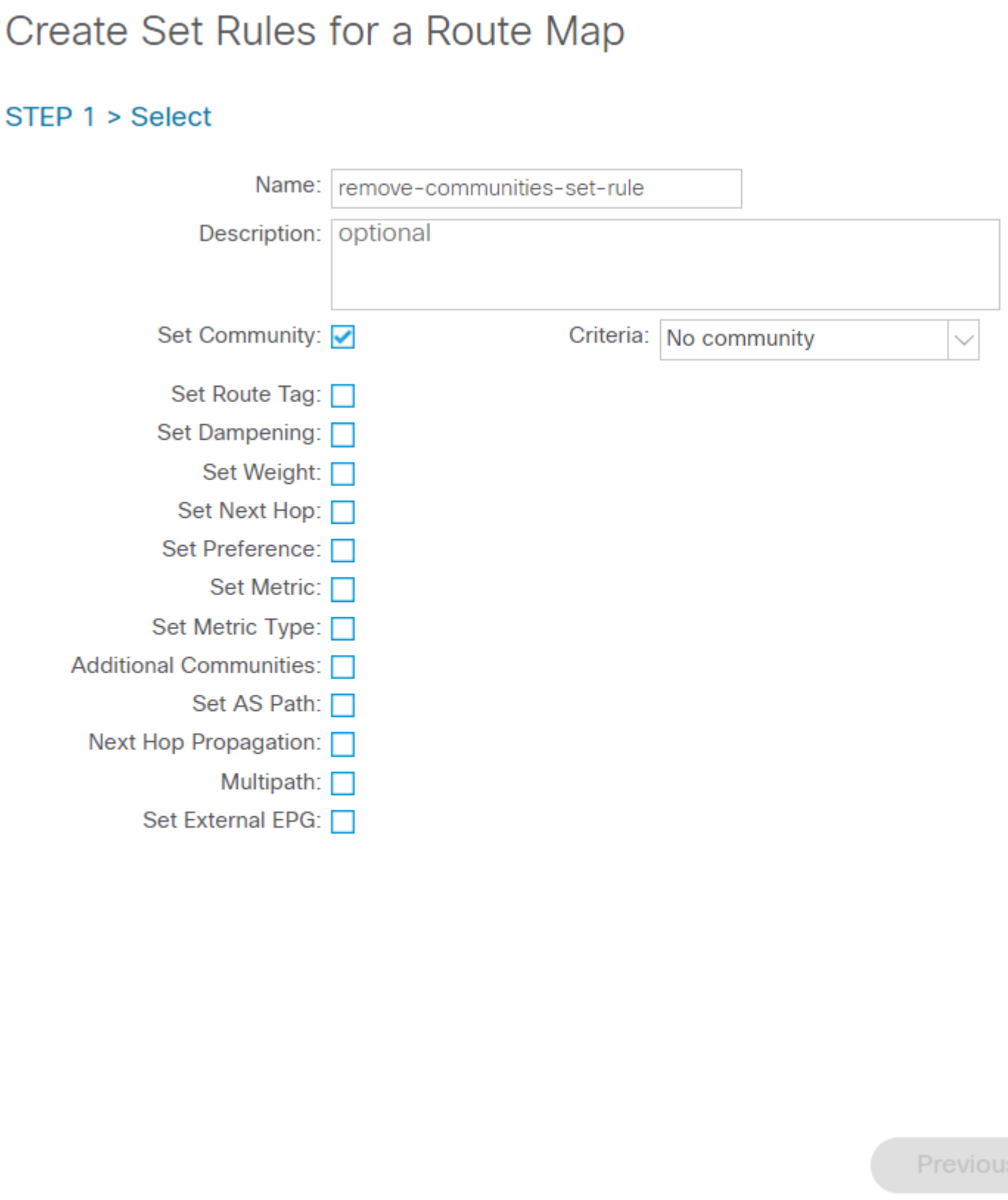
La regla de coincidencia asociada ahora se agrega al contexto de control de ruta

A continuación, seleccione el menú desplegable junto a Set Rule y seleccione Create Set Rules for a Route Map:



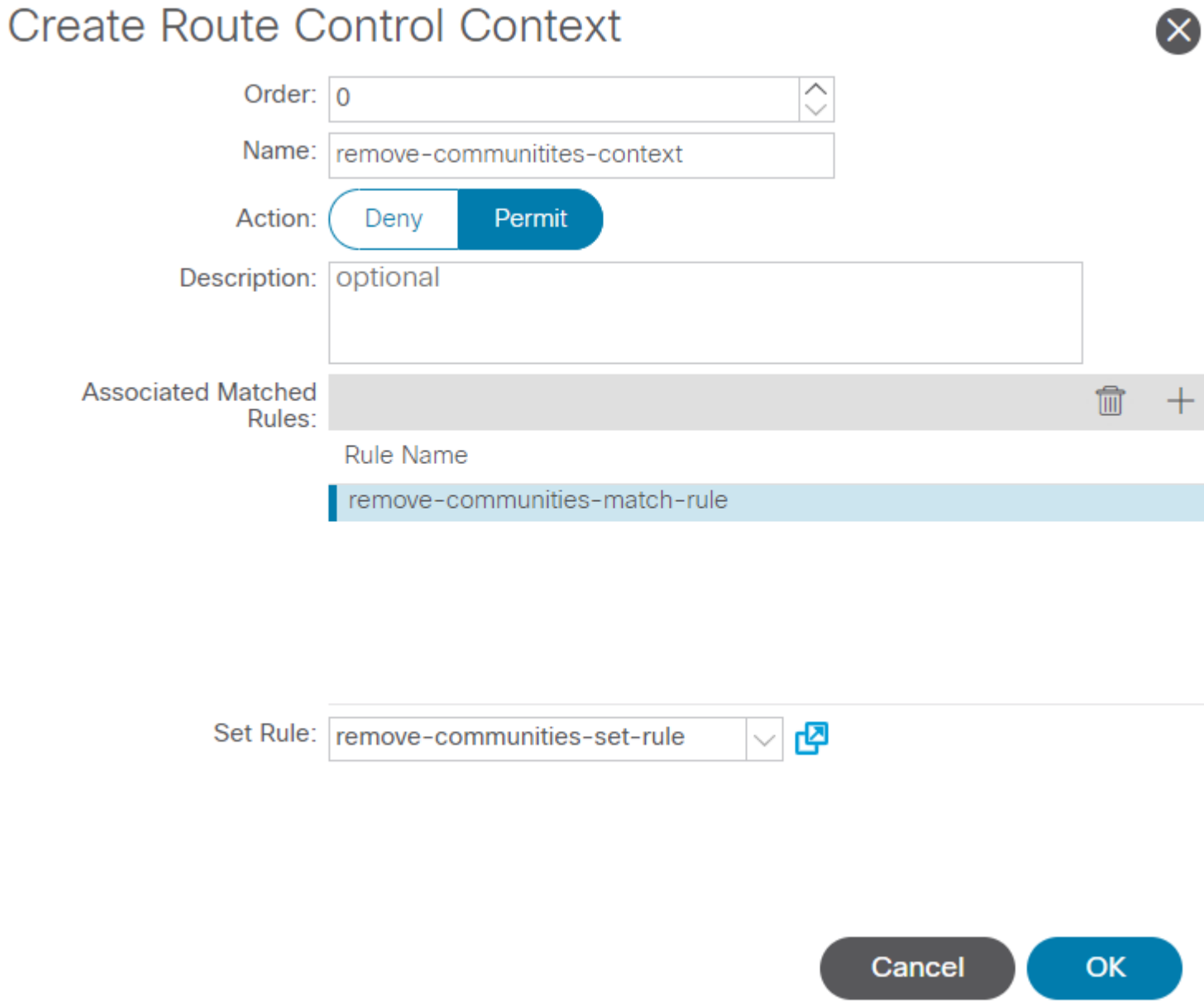
Seleccione esta opción para crear reglas de conjunto para un mapa de ruta

Asigne un nombre a la regla establecida y, a continuación, seleccione la Set Community y deje los criterios predeterminados de No community seleccionado:



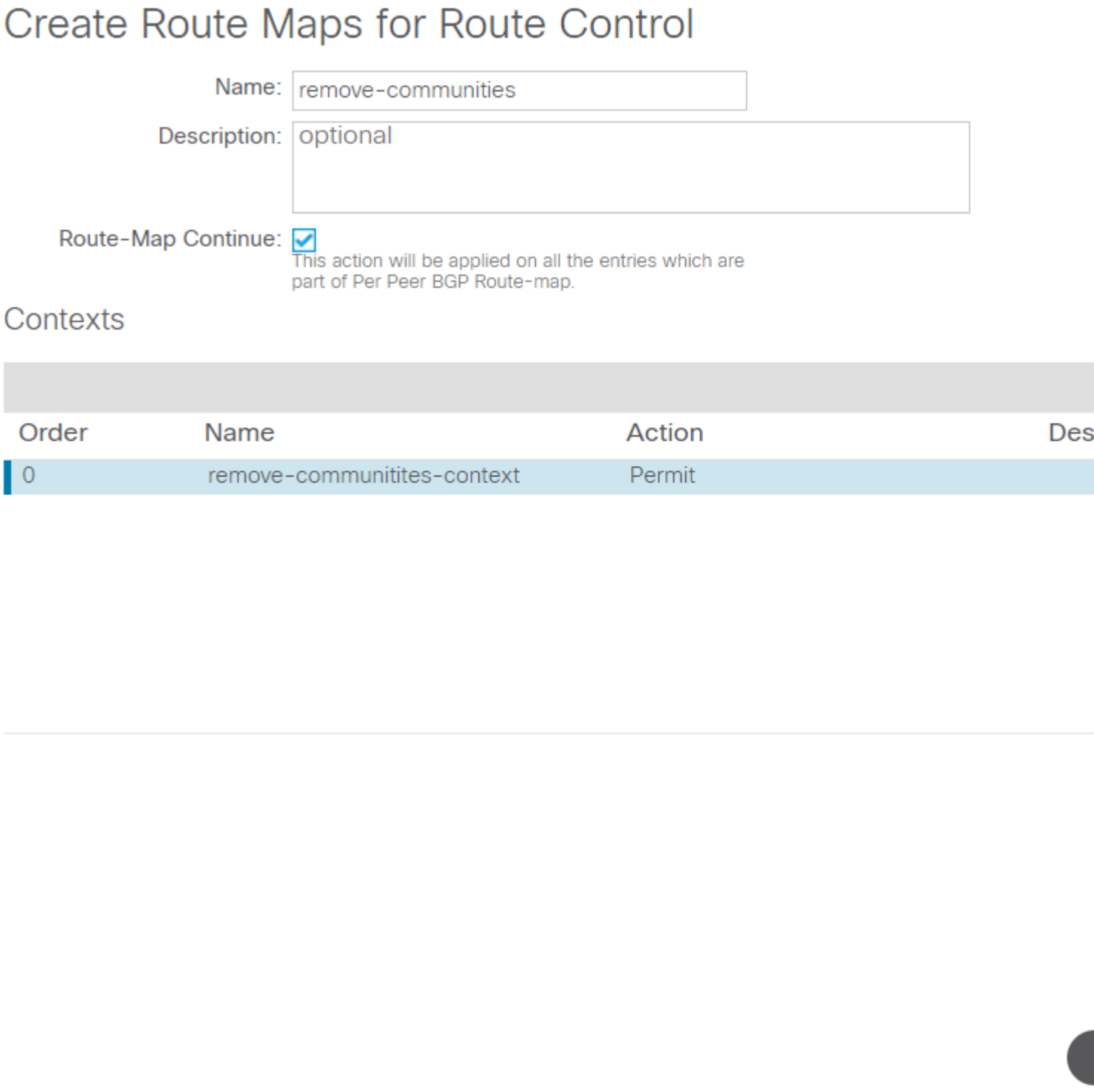
Crear regla de conjunto para mapa de ruta

Después de seleccionar Finalizar en el Create Set Rules for a Route Map , verá la regla definida seleccionada en la ventana Create Route Control Context ventana:



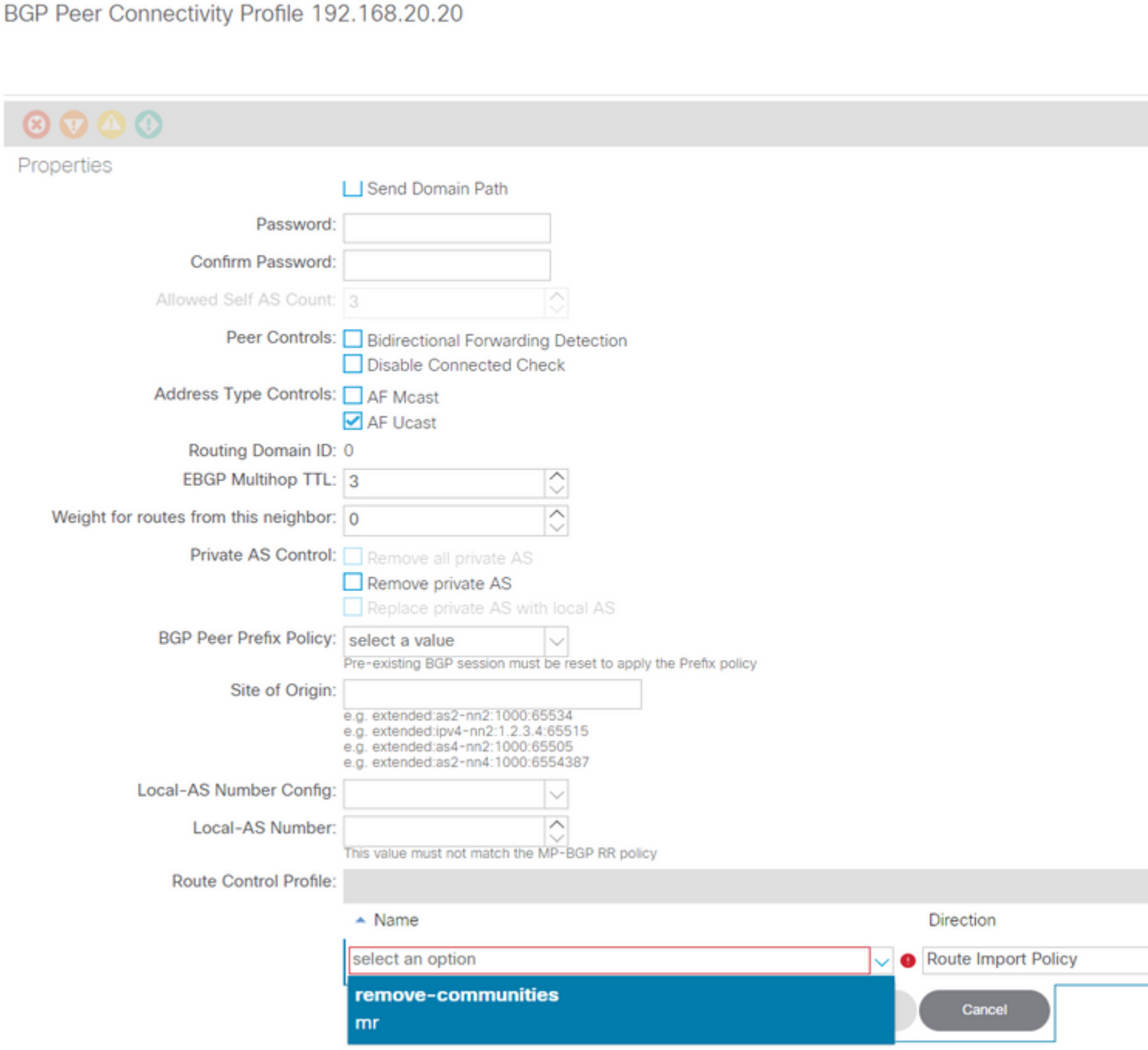
La regla de conjunto se agrega ahora al contexto de control de ruta

Después de seleccionar OK en el Create Route Control Context , verá el contexto agregado a la ventana Contexts tabla en el Create Route Maps for Route Control ventana. Por último, seleccione Submit para completar la configuración:



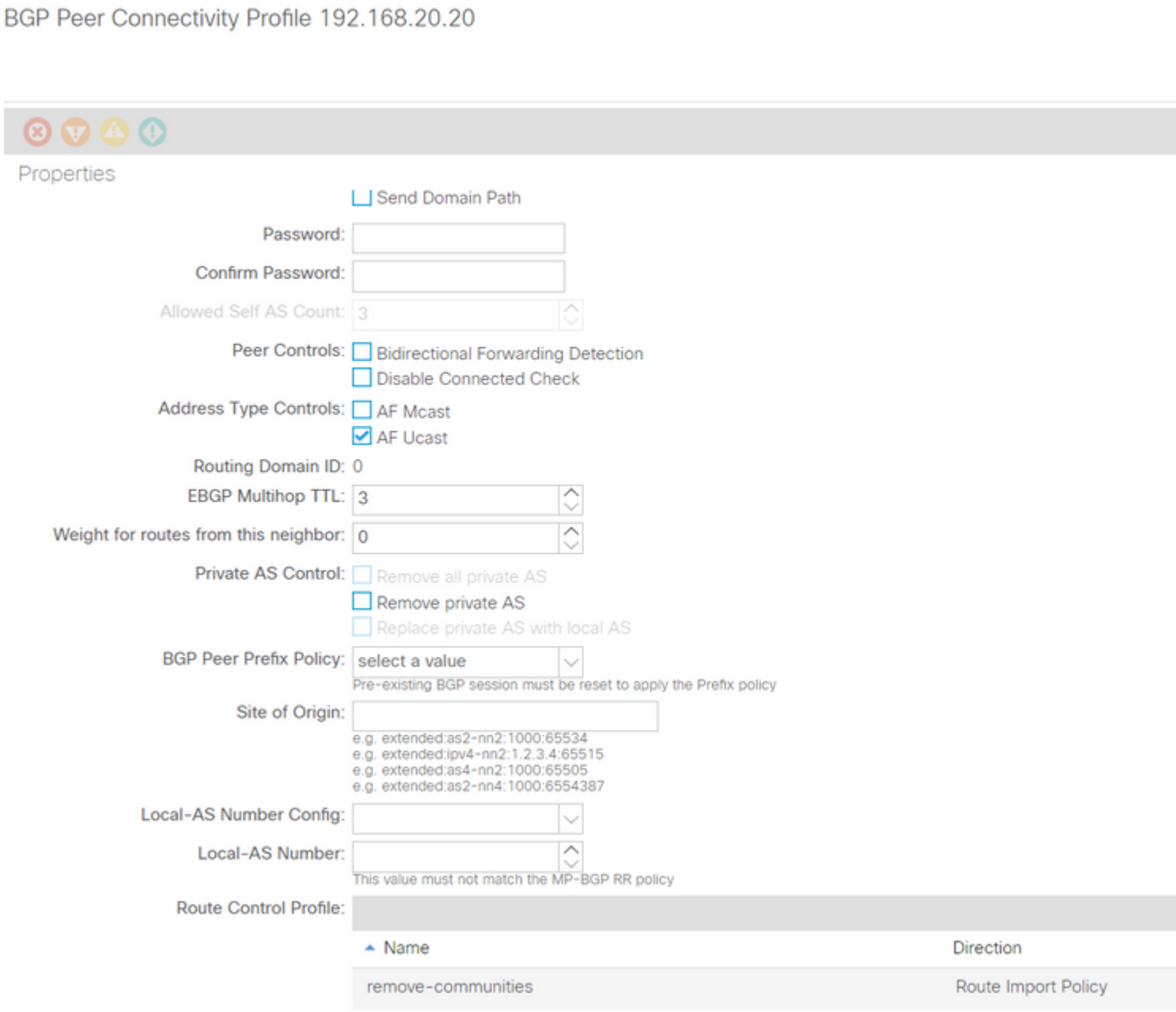
El contexto se agrega ahora al mapa de ruta

Navegue hasta el perfil de conectividad de par BGP en L3Out y seleccione el + icono en el Route Control Profile y, a continuación, agregue el mapa de ruta con la dirección predeterminada de Route Import Policy seleccionado:



Agregar mapa de ruta al perfil de conectividad de par BGP

Después de seleccionar **Update** para el route map, verá su route map agregado al Route Control Profile tabla:



El mapa de ruta ahora se agrega al perfil de conectividad de par BGP

*Para obtener más información sobre las opciones de configuración de route map en ACI, consulte el [informe técnico de ACI Fabric L3Out]

Después de implementar una de las soluciones anteriores, verifique si el problema está resuelto.

Verifique el estado de BGP en BL 101:

<#root>

```
leaf-101# show ip bgp 192.168.20.0 vrf example:example
BGP routing table information for VRF example:example, address family IPv4 Unicast
BGP routing table entry for 192.168.20.0/24, version 46 dest ptr 0xa0fec840
Paths: (1 available, best #1)
Flags: (0x80c001a 00000000) on xmit-list, is in urib, is best urib route, is in HW, exported
  vpn: version 2731, (0x100002) on xmit-list
Multipath: eBGP iBGP
```

```
  Advertised path-id 1, VPN AF advertised path-id 1
  Path type (0xa96485b8): internal 0x18 0x0 ref 0 adv path ref 2, path is valid, is best path
  AS-Path: NONE, path sourced internal to AS
    192.168.20.20 (metric 5) from 192.168.20.20 (192.168.20.20)
      Origin IGP, MED not set, localpref 100, weight 0 tag 0, propagate 0
      Extcommunity:
          RT:65001:2162688
          COST:pre-bestpath:163:1879048192


***Notice that no router mac is present here.***


          VNID:2162688

  VRF advertise information:
  Path-id 1 not advertised to any peer

  VPN AF advertise information:
  Path-id 1 advertised to peers:
    10.0.216.65        10.0.216.66
```

Compruebe el NERVIO en CL 102:

<#root>

```
leaf-102# show ip route 192.168.20.0 vrf example:example
IP Route Table for VRF "example:example"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

192.168.20.0/24, ubest/mbest: 1/0
    *via 10.0.210.70%overlay-1, [200/0], 00:00:06, bgp-65001, internal, tag 65001
         recursive next hop: 10.0.210.70/32%overlay-1


***Notice that no rwVnid entry is present here.***
```

> **Nota**: La ausencia o presencia de la entrada rwVind por sí sola no determina si el problema se está produciendo o no. En muchos casos, la entrada rwVind se elimina de la ruta en cuestión una vez que se resuelve el problema. Sin embargo, no siempre es así. Siempre verifique las tablas FIB y HAL para verificar si el problema se resuelve o no.

Comprobar FIB en CL 102:

<#root>

```
module-1(DBG-elam-insel6)# show forwarding route 192.168.20.0 vrf example:example

IPv4 routes for table example:example/base

------------------+------------------+----------------------+------------------------
```

```
Prefix            | Next-hop         | Interface/VRF        | Additional Info
------------------+------------------+----------------------+------------------------
```

***192.168.20.0/24**

10.0.210.70

overlay-1

******Notice that we have the route here and our next-hop address is correct (showing the TEP IP of BL 101)***

```
Route Class-id:0x0
Policy Prefix 0.0.0.0/0

leaf-102# acidiag fnvread | grep 101
     101        1             leaf-101
```

10.0.210.70/32

```
    leaf         active   0
```

Tabla HAL en CL 102:

<#root>

```
module-1(DBG-elam-insel6)# show platform internal hal l3 routes | grep 192.168.20.0
|
```

4662

```
| 192.168.20.0/ 24| UC|  686|     20601| TRIE|   a5| 5/  0| 60a5|A|     8443|    86b6|  ef5| 1/  2|
```

******Notice that we have an entry here and it's in the correct VRF.******

```
module-1(DBG-elam-insel6)# hex
```

4662

```
0x
```

1236

```
module-1(DBG-elam-insel6)# show platform internal hal l3 vrf pi
=============================================================================================================
                                    | -- TOR --  | - Spine -  |                ACL                 |
      Vrf           Hw    I I Vrf     | SB    NB   | Proxy ACI  |      Ing              Egr          |
VrfId Name          VrfId I S Vnid    | BDId  BDId | Ou Bd Enc  | Lbl      Msk      Lbl      Msk      |
=============================================================================================================
26    example:example
```

1236

```
0 0 210000    0     0      0     1      0        0        0        0        0
```

Pings de EP (Host 1) al host en una red externa que proviene de un peer BGP externo (192.168.20.20):

<#root>

```
Host-1# ping 192.168.20.20 vrf example
PING 192.168.20.20 (192.168.20.20): 56 data bytes
64 bytes from 192.168.20.20: icmp_seq=0 ttl=252 time=1.043 ms
64 bytes from 192.168.20.20: icmp_seq=1 ttl=252 time=1.292 ms
64 bytes from 192.168.20.20: icmp_seq=2 ttl=252 time=1.004 ms
64 bytes from 192.168.20.20: icmp_seq=3 ttl=252 time=0.769 ms
64 bytes from 192.168.20.20: icmp_seq=4 ttl=252 time=1.265 ms

--- 192.168.20.20 ping statistics ---
5 packets transmitted, 5 packets received, 0.00% packet loss
round-trip min/avg/max = 0.769/1.074/1.292 ms
```

******Connectivity is there.******

ELAM en CL 102:

<#root>

```
leaf-102# vsh_lc
module-1# debug platform internal roc elam asic 0
module-1(DBG-elam)# trigger reset
module-1(DBG-elam)# trigger init in-select 6 out-select 0
module-1(DBG-elam-insel6)# set outer ipv4 src_ip 192.168.10.10 dst_ip 192.168.20.20
module-1(DBG-elam-insel6)# start
module-1(DBG-elam-insel6)# stat
 ELAM STATUS
===========
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Triggered

module-1(DBG-elam-insel6)# ereport
Python available. Continue ELAM decode with LC Pkg
 ELAM REPORT
<output omitted>
--------------------------------------------------------------------------------------------------------------
Lookup Drop
--------------------------------------------------------------------------------------------------------------
LU drop reason                          :
```

no drop

******Traffic forwards correctly.******

Información Relacionada

- Este comportamiento también se documenta en este defecto: ID de bug de Cisco CSCvx28929
- **Soporte Técnico y Documentación - Cisco Systems**

Acerca de esta traducción
Cisco ha traducido este documento combinando la traducción automática y los recursos humanos a fin de ofrecer a nuestros usuarios en todo el mundo contenido en su propio idioma.

Tenga en cuenta que incluso la mejor traducción automática podría no ser tan precisa como la proporcionada por un traductor profesional.

Cisco Systems, Inc. no asume ninguna responsabilidad por la precisión de estas traducciones y recomienda remitirse siempre al documento original escrito en inglés (insertar vínculo URL).