EX ハードウェア:ACI パケット転送の詳細

内容

概要 前提条件 要件 使用するコンポーネント シナリオ 同じ EPG/同じリーフでの2つの EP - スイッチド フレーム トポロジ **ELAM** <u>別の EPG/同じリーフでの 2 つの EP - ルーテッド パケット</u> トポロジ **ELAM** <u>別の EPG/別のリーフでの 2 つの EP - ルーテッド パケット</u> トポロジ **ELAM** 1 EP ---> L3アウト - ルーテッドフロー トポロジ **ELAM** 1EP —>リモートEPまたはSVI – スパイン検証 トポロジ ロジック シンセティック IP ファブリック モジュール ELAM 追加シナリオ:「hal internal-port pi」出力にない Ovector の取得 トポロジ ロジック

概要

このドキュメントでは、アプリケーションセントリックインフラストラクチャ(ACI)の「EX」ベ ースのACIスイッチを使用する異なる転送シナリオについて説明します。ハードウェアが正しく プログラムされ、適切なエンドポイントグループ(EPG)の正しい宛先エンドポイント(EP)にパケ ットを転送します。

前提条件

要件

このドキュメントに特有の要件はありません。

使用するコンポーネント

このドキュメントの情報は、次のハードウェアとソフトウェアのバージョンに基づいています。

- EX ハードウェアを使用する、スパイン スイッチと2つのリーフ スイッチによって構成される ACI ファブリック
- 各リーフ スイッチに向かう 2 つのアップリンクを持つ ESXi ホスト
- ルータとして機能する Nexus 5000 デバイス
- 初期設定に使用されるアプリケーション ポリシー インフラストラクチャ コントローラ (APIC)

このドキュメントの情報は、特定のラボ環境にあるデバイスに基づいて作成されました。このド キュメントで使用するすべてのデバイスは、初期(デフォルト)設定の状態から起動しています 。対象のネットワークが実稼働中である場合には、どのようなコマンドについても、その潜在的 な影響について確実に理解しておく必要があります。

シナリオ

同じ EPG/同じリーフでの 2 つの EP - スイッチド フレーム

トポロジ

0050.56a5.fccc

192.168.20.2/24



EPG1 0050.56a5.6794 192.168.20.3/24

このトポロジの場合、EP1 から EP2 へのフローは L2 フローになり、送信元トラフィックがどの リーフに入ってくるとしても、ローカルで切り替える必要があります。 レイヤ 2(L2)のフロー について、MAC アドレス テーブルを最初に確認し、スイッチがフレームを受信しているか、ま たどこで受信しているかを判別する必要があります。

leaf4# show mac address-table | grep fccc
* 30 0050.56a5.fccc dynamic - F F po3
leaf4# show mac address-table | grep 6794
* 30 0050.56a5.6794 dynamic - F F po4
カプセル化 VLAN を確認するために、EP データベースを確かめることもできます。

leaf4# show endpoint mac 0050.56a5.fccc Legend: 0 - peer-attached H - vtep a - locally-aged S - static L - local V - vpc-attached p - peer-aged M - span B - bounce s – static-arp ---+ MAC Address VLAN/ MAC Info/ Encap Interface IP Address Domain VLAN IP Info ---+ vlan-2268 0050.56a5.fccc LV 30 po3 Joey-Tenant:Joey-Internal vlan-2268 192.168.20.2 LV po3 calo2-leaf4# show endpoint mac 0050.56a5.6794 Legend: 0 - peer-attached H - vtep a - locally-aged S - static V - vpc-attached p - peer-aged L - local M - span B - bounce s - static-arp VLAN/ Encap MAC Address MAC Info/ Interface VLAN IP Address IP Info Domain ---+ 30 vlan-2268 0050.56a5.6794 LV po4 vlan-2268 192.168.20.3 LV Joey-Tenant: Joey-Internal po4

FD_VLAN 30 が一致していることが分かりますが、いつでもソフトウェアでのマッピングを検証 することができます。

leaf4# show vlan extended | grep 2268

30 enet CE vlan-2268

もちろん、前面パネルのカプセル化として VLAN 30 が VLAN 2268 にマッピングされることを確認するためにハードウェアを確かめることができます。

leaf4# vsh_lc
module-1# show system internal eltmc info vlan 30

vlan_id:	30	:::	hw_vlan_id:	22
vlan_type:	FD_VLAN	:::	bd_vlan:	28
access_encap_type:	802.1q	:::	access_encap:	2268
fabric_encap_type:	VXLAN	:::	fabric_encap:	11960
sclass:	32778	:::	scope:	11
untagged:	0			
acess_encap_hex:	0x8dc	:::	fabric_enc_hex:	0x2eb8
pd_vlan_ft_mask:	0x8			
fd_learn_disable:	0			
qos_class_id:	0	:::	qos_pap_id:	0
qq_met_ptr:	25	:::	<pre>ipmc_index:</pre>	0
ingressBdAclLabel:	0	:::	ingBdAclLblMask:	0
egressBdAclLabel:	0	:::	eqrBdAclLblMask:	0

qos_map_idx:	0	:::	qos_map_pri:	0
qos_map_dscp:	0	:::	qos_map_tc:	0
vlan_ft_mask:	0xe30			
hw_bd_idx:	0	:::	hw_epg_idx:	11267
intf_count:	2	:::	glbl_scp_if_cnt:	2

<SNIPPED>

EP がソフトウェアで学習されることを考えると、ハードウェアでこれらの EP の L2 情報がプロ グラムされていることを確認することもできます。 新しいハードウェアには、ハードウェアにお けるソフトウェアの状態を示すハードウェア アブストラクション レイヤ(HAL)があります。 HAL の役割は、ソフトウェア プログラミング要求を受け取って、ハードウェアに渡すことです 。

エンドポイントに関する L2 ハードウェア情報を確認するには、特定の MAC アドレスの HAL の L2 テーブルを参照できます。

<pre>leaf4# vsh_lc module-1# show pla LEGEND:</pre>	atform internal ha	Lep 12 ma	ac 0050.56a5	.fccc					
BDId:	BD Id					BD	Name:		BD
Name									
т:	EP Type (Pl: Phys	sical Vl:	Virtual Xr:	Remot	ce	EP	Mac:		Mac
L2 IfId:	L2 Interface					L2	IfName	:	L2
IfName									
FDId:	FD Id					FD	Name:		FD
Name									
S Class:	S Class					Aqe	e Intvl	:	Aqe
Interval						2			5
P A:	Packet Action (F	Forward	, T: Trap to	CPU,					
	L	Log & Fa	orward, D: D:	, 1, qor	N: None)				
S T:	Static Ep	2		1,	,	SI	Ξ:		
Secure EP	-								
L D:	Learn Disable					ві	N D:		Bind
Notify Disable									
E N D:	Epg Notify Disabl	le				B	E :		
Bounce Enable									
I D L:	IVxlan Dont Learn	ı				SP	Ι:		
Source Policy Inc	omplete								
DPI:	Dest Policy Incor	mplete				SP	A:		
Source Policy App	lied								
DPA:	Dest Policy Appl	ied				DS	s:		Dest
Shared Service									
IL:	Is Local					VUI	в:		Vnid
Use Bd									
so:	SA Only								
L2 EP Count: 1									
	=====								
ISDSDD V									ВЕ
BD	EP	L2	L2		FD	S	Age	PSS	L N N
B D P P P S I U	S								
BdId Name T	Mac	IfId	Ifname	FDId	Name	Class	Intvl	ΑΤΕ	DDD
E L I I A A S L B	0					.====:		.====:	
	=====	16000000	D o ²	10	ED 30	000-	205		0 1 0
	OUIDUIDOIADIICICC	T000005	P03	те	FD-20	auua	29 <u>1</u>	r U U	UIU
	U								

module-1# show platform internal hal ep 12 mac 0050.56a5.6794 _____ ΒE ISDSDD V BD L2 ΕP T.2 FD S Age PSSLNN BDPPPPSIUS BdId Name T Mac TfTd Ifname FDId Name Class Intvl A T E D D D ELIIAASLBO _____ _____ Pl 00:50:56:a5:67:94 16000003 Po4 1e **FD-30** 800a 29f F 0 0 0 1 0 1cBD-28 0 0 0 0 0 0 0 1 0 0 ここまでで、ハードウェアについてまとめることができましたが、今度は ELAM について、また パケットがどこに送信される必要があるかを考えましょう。

ELAM

module-1(DBG-TAH-elam-insel6)# report | grep ovec sug_elam_out_sidebnd_no_spare_vec.ovector_idx: 0x9E

module-1(DBG-TAH-elam-insel6)# show platform internal hal 12 port gpd

Leaf4はAsic 0スライス1のフレームを受信しました。新しいハードウェアのELAMでは、トラブ ルシューティングに非常に重要な新しいフィールドがあります。ovector_idx があります。 この インデックスは、フレーム/パケットが転送される物理ポートのインデックスです。 ovector_idx を設定すると、このコマンドを使用して、どのポートにマッピングされているか調べることがで きます。

Legend: _____ IfId: Interface Id IfName: Interface Name Is PC Mbr IfId: I P: Interface Id Uc PC Cfg: UcPcCfg Idx Uc PC MbrId: Uc Pc Mbr Id Asic Port Asic AP: As: Sl: Slice Sp: Slice Port Ovector (slice | Ss: Slice SrcId Ovec: srcid) L S: Local Slot Reprogram: L3: Is L3 P: PifTable Xla Idx: Xlate Idx Rw PifTable Ovx Idx: OXlate Idx RP: IP: If Profile Table N L3: Num. of L3 Ifs RS: Rw SrcId Table NI L3: Num. of Infra L3 Ifs DP: DPort Table Vif Tid: Vif Tid SP: SrcPortState Table RwV Tid: RwVif Tid

RSP: RwSrcPortstate Table Ing Lbl: Ingress Acl Label UC: UCPcCfg Egr Lbl: Egress Acl Label UM: UCPcMbr Reprogram: PROF ID: Lport Profile Id VS: VifStateTable HI: LportProfile Hw Install RV: Rw VifTable Num. of Sandboxes: 1 Sandbox_ID: 0, BMP: 0x0 Port Count: 8 _____ UC UC Reprogram Rep I PC Pc L | RIRD R UUX | L Xla Ovx N NI Vif RwV Ing Egr | V R | PROF H IfId Ifname P Cfg MbrID As AP Sl Sp Ss Ovec S | P P P S P Sp Sp C M L | 3 Idx Idx L3 L3 Tid Tid Lbl Lbl | SV | ID I _____ _____ 1a004000 Eth1/5 1 0 1d 0 d 0 c 18 18 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 1a005000 Eth1/6 10 b 0 e 0 d 1a 1a 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 02650f0e1c1c10000000000 1a006000 Eth1/7 0 0 0 0 0 D-256 - 800 0 0 1 e 0 0 10 0 f 1e 1e 1 0 0 0 0 0 0 0 0 0 0 0 0 02e 7 0 0 1a007000 Eth1/8 D-84 - 800 0 0 1 30 0 1a01e000 Eth1/31 1 0 2d 0 37 1 e 1c 9c 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 0 0 010 0 **1a01f000 Eth1/32** 1 0 3d 0 38 1 f le **9e** 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - 0 0 0 1 0 0 1a030000 Eth1/49 0 2 1 0 49 1 20 38 b8 1 0 0 0 0 0 0 0 0 0 1 8 6 2 2 D-24d - 400 0 0 0 1 0 0 29 1 0 0 80 1 0 0 0 0 0 0 0 0 0 1 9 7 2 2 1a031000 Eth1/50 03 3 D-350 - 400 0 0 0 1 0 スイッチは、パケットがインターフェイスEthernet 1/32から転送されると考えています。MACア

ドレスを学習したPO4ですか。

leaf4‡	\$ show	port-channe	el summary	
Flags:	D -	Down	P - Up in po	rt-channel (members)
	I -	Individual	H - Hot-stan	dby (LACP only)
	s -	Suspended	r - Module-r	emoved
	S -	Switched	R - Routed	
	U -	Up (port-ch	annel)	
	м –	Not in use.	Min-links no	t met
	F -	Configurati	on failed	
Group	Port-	Туре	Protocol	Member Ports
	Channe	el		
1	Pol(St	J) Eth	LACP	Eth1/5(P)
2	Po2(St	J) Eth	LACP	Eth1/6(P)
3	Po3(St	J) Eth	LACP	Eth1/31(P)
4	Po4 (St	J) Eth	LACP	Eth1/32(P)
そのと	こおり	です。パケ	ットはインタ	ーフェイス 1/32 から宛先ホストに転送されます。

別の EPG/同じリーフでの 2 つの EP - ルーテッド パケット



この例では、EP1 から EP2 へのパケットのパケット フローを追跡します。EP1 と EP2 は同じ vPC リーフ ペアに存在しています。 2 つの EP は、異なる BD を使用する、異なる EPG に存在 します。

最初に EP データベースを確認し、EP を学習しているかどうかを確かめます。

leaf4# show endpoint	ip 192.168.20.2				
0 - peer-attached	H - vtep	a - locali	Ly-aged S -	static	
V - vpc-attached	p – peer-aged	L - local	M –	span	
s - static-arp	B - bounce				
+	+	++	+	+	+
VLAN/		Encap	MAC Address	MAC Info/	
Interface					
Domain		VLAN	IP Address	IP Info	
++ 30	4	vlan-2268	0050.56a5.fcc	c LV	
po3 Joev-Tenant:Joev-Inte	rnal	vlan-2268	192.168.20.1	2 T.V	
po3					
calo2-leaf4# show end	lpoint ip 192.168	3.21.2			
	TT and an	- level]	a and a		
0 - peer-attached	н – vlep	a - locali	Ly-aged S - :	SLALIC	
V - vpc-attached	p - peer-aged	L - local	M –	span	
s – static-arp	B – bounce				

+	+	+	-+
+			
VLAN/	Encap	MAC Address	MAC Info/
Interface			,
Domain	VLAN	IP Address	IP Info
+	+	+	-+
+			
8	vlan-2200	0050.56a5.0c11	LV
po4			
Joey-Tenant:Joey-Internal	vlan-2200	192.168.21.2	LV
po4			

EP はすでに学習されており、IP 情報を把握しているため、ハードウェアの EP 学習情報を確認 することができます。

<pre>leaf4# vsh_lc module-1# show plat LECEND:</pre>	tform internal hal ep 13 all														
LEGEND.															
VrfName:	Vrf Name								т:						Type
(Pl: Physical, Vl:	Virtual, Xr: Remote)														11 -
EP IP:	Endpoint IP														
S Class:	S Class								Ag	еI	ntv	71:			Age
Interval															
S T:	Static Ep								S	Е:					
Secure EP															
L D:	Learn Disable								в	N D	:				Bind
Notify Disable															
E N D:	Epg Notify Disable								В	E:					
Bounce Enable															
I D L:	IVxlan Dont Learn								SP	1:					
Source Policy Incom	mplete														
DPI:	Dest Policy Incomplete								SP.	A:					
Source Policy Appl:	ied														
DPA:	Dest Policy Applied								DS	s:					Dest
Shared Service															
IT:	Is Local								VU	в:					Vnid
Use Bd															
SO:	SA Only								ΕP	NH	L3	BIfN	Jam	ie:	EP
Next Hop L3 If Name	e														
NHT:	Next Hop Type (L2: L2 Entry	L3: L3	Next H	lop)					BD	Na	me:				L2 NH
BD Name															
EP Mac:	EP Mac								Г3	If	Nan	ne:			L3 NH
If Name															
L2 IfName:	L2 If Name								FD	Na	me:				L2
Entry FD Name															
ID:	L3 NH IP														
L3 EP Count: 12															
			======	====	===	===	==:	===	===	===	===	===	===	===	
						в	E –	 I	s	DS	D	D	v	,	- EP-NH
N															
Vrf El	P	S	Age	S	S L	N I	NI	ВD	Ρ	ΡP	Ρ	S I	L U	S	L3
H BD EP	L3	L2	FI)											
Name T II	P	Clas	s Intvl	LТ	E D	D	DI	ΕL	Ι	ΙA	Α	SΙ	B	0	
IfName T 1	Name Mac	IfName	I	Efna	ame]	Nam	е		ΙF)			
		======	======			===	==:	===	===	===	===		===	===	
		======	======	====	===	===	==:	===	===	===	===		===	===	-
common*rewall Pl 1	0.6.112.1	1	0	1	0 0	0	0	0 1	1	0 0	0	0 1	L 0	0	-

L3	-	00	00:00:00:00:00	-	-			-				0	.0.	0.	0								
commo	on*rewall	Pl	10.6.114.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	0.	0								
commo	on*rewall	Pl	10.6.114.129			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								
commo	n*efault	Pl	100.100.101.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								
Joey-	T*ternal	Pl	192.168.1.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
LЗ	-	00	0:00:00:00:00:00	-	-			-				0	.0.	0.	0								
Joey-	-T*ternal	Xr	192.168.1.100			8013	128		0	0	0	1 (0 0	0	0	0	0	0	0	0	1	0	-
L3	-	00):0c:0c:0c:0c:0c	Tunnel2	Tur	nnel2		-				0	.0.	Ο.	0								
Joey-	T*ernal2	Pl	192.168.3.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								
Joey-	T*ternal	Pl	192.168.20.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								
Joey-	T*ternal	P1	192.168.20.2			800a	0		0	0	0	0 0	0 0	0	0	0	0	0	0	1	0	0	-
L2	BD-28	00):50:56:a5:fc:cc	-	Po	3		FD-	30)		-											
Joey-	-T*ternal	Pl	192.168.21.1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								
Joey-	T*ternal	P1	192.168.21.2			800c	0		0	0	0	0 0	0 0	0	0	0	0	0	0	1	0	0	-
L2	BD-7	00):50:56:a5:0c:11	-	Poé	1		FD-	8			-											
Joey-	T*ternal	Pl	2001:0:0:100::1			1	0		1	0	0	0 0	0 0	1	1	0	0	0	0	1	0	0	-
L3	-	00	0:00:00:00:00:00	-	-			-				0	.0.	Ο.	0								

HAL Layer3(I3)のテーブルでは、I3 学習済み EP に関する VLAN/ポート情報が提供されるため 非常に役立ちます。 Po4 の宛先が存在するため、パケットは Po4 のいずれかのポートから転送 される必要があります。

ELAM を実行してみて、どのような結果が得られるか確認してみましょう。

ELAM

Asic 0 Slice 1 Status Triggered

module-1(DBG-TAH-elam-insel6)# report | grep ovec sug_elam_out_sidebnd_no_spare_vec.ovector_idx: 0x9E

パケットをトリガーし、「ovector_idx」が0x9Eであることがわかりました。ovectorインデック スは、パケットを転送する必要がある発信の物理インターフェイスインデックスです。 どのポー トがそのインデックスに設定されているか見てみましょう。

IfId: Interface Id IfName: Interface Name Is PC Mbr IfId: Interface Id T P: Uc PC Cfg: UcPcCfg Idx Uc PC MbrId: Uc Pc Mbr Id Asic Port As: Asic AP: Sl: Slice Sp: Slice Port Slice SrcId Ovec: Ovector (slice | Ss: srcid) Local Slot L S: Reprogram: г3: Is L3 P: PifTable Xla Idx: Xlate Idx RP: Rw PifTable Ovx Idx: OXlate Idx IP: If Profile Table N 1.3: Num. of L3 Ifs Num. of Infra L3 Ifs RS: Rw SrcId Table NT L3: DP: DPort Table Vif Tid: Vif Tid SP: SrcPortState Table RwV Tid: RwVif Tid RSP: RwSrcPortstate Table Ing Lbl: Ingress Acl Label Egr Lbl: UC: UCPcCfg Egress Acl Label UM: UCPcMbr Reprogram: PROF ID: Lport Profile Id VS: VifStateTable HI: LportProfile Hw Install RV: Rw VifTable Num. of Sandboxes: 1 Sandbox_ID: 0, BMP: 0x0 Port Count: 8 _____ _____ UC UC Reprogram Rep I PC Pc L | RIRD R UUX | L Xla Ovx N NI Vif RwV Ing Egr | V R | PROF H IfId Ifname P Cfg MbrID As AP Sl Sp Ss Ovec S | P P P S P Sp Sp C M L | 3 Idx Idx L3 L3 Tid Tid Lbl Lbl | SV | ID I _____ -----1a004000 Eth1/5 1 0 1d 0 d 0 c 18 18 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 1a005000 Eth1/6 10 b 0 e 0 d 1a 1a 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 1a006000 Eth1/7 02650f0e1c1c1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 D-256 - 800 0 0 1 c 0 0 2f 7 0 10 0 f le le 0 0 0 0 0 0 0 0 0 0 1a007000 Eth1/8 1 0 0 0 0 0 0 1 D-199 - 800 0 2e 0 1a01e000 Eth1/31 1 0 2d 0 37 1 e 1c 9c 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 0 0 0 1 0 0 1a01f000 Eth1/32 1 0 3d 0 38 1 f 1e 9e 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 - 0 0 01 0 0 1a030000 Eth1/49 0 2 1 0 49 1 20 38 b8 1 0 0 0 0 0 0 0 0 0 16 4 2 2 D-24d - 400 0 0 0 1 0
 la031000 Eth1/50
 0 3
 3
 0
 29 1
 0
 80
 1
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0 15 3 2 2 D-350 - 400 0 0 0 1 0

ポート 1/32 から送信する必要があるようですが、これは正しいでしょうか。

leaf4# show port-channel summary
Flags: D - Down P - Up in port-channel (members)
I - Individual H - Hot-standby (LACP only)
s - Suspended r - Module-removed
S - Switched R - Routed
U - Up (port-channel)
M - Not in use. Min-links not met

F -	Configuration	failed
-----	---------------	--------

Group	Port- Channel	Туре	Protocol	Member Ports
1	Pol(SU)	Eth	LACP	Eth1/5(P)
2	Po2(SU)	Eth	LACP	Eth1/6(P)
3	Po3(SU)	Eth	LACP	Eth1/31(P)
4	Po4(SU)	Eth	LACP	Eth1/32(P)
はい、	これは適切]です。		

別の EPG/別のリーフでの 2 つの EP - ルーテッド パケット

トポロジ



この例では、EP1 から EP2 へのパケットのパケット フローを追跡します。EP1 は EX vPC ペア に存在し、EP2 はリモートの Generation 1 vPC リーフのペアに存在しています。 2 つの EP は 、異なる BD を使用する、異なる EPG に存在します。

もう一度、EP が学習されている場所を確認してみましょう。

<pre>leaf4# show endpoint Leagued:</pre>	ip 192.168.20.2				
0 - peer-attached V - vpc-attached s - static-arp	H - vtep p - peer-aged B - bounce	a - local L - local	ly-aged S - M -	static span	
++		+	+	+	-+
VLAN/		Encap	MAC Address	MAC Info/	
Interface					
Domain		VLAN	IP Address	IP Info	
+		+	+	+	-+
+		w1an_2268	0050 5625 fa		
20		v 1411-2200	0050.5045.10		
Joey-Tenant:Joey-Inte	ernal	vlan-2268	192.168.20	.2 LV	

calo2-leaf4# show en	dpoint ip 192.168	3.1.100			
Legend:					
0 - peer-attached V - vpc-attached s - static-arp	H – vtep p – peer-aged B – bounce	a - local L - local	ly-aged S - M -	- static - span	
++		+	+	++-	
VLAN/		Encap	MAC Address	MAC Info/	
Interface					
Domain		VLAN	IP Address	IP Info	
+		+	+	++-	
+					
Joey-Tenant:Joey-Int	ernal		192.168.1.1	00	
. 10					

```
Joey-Tena
tunnel2
```

ここで、ハードウェアがどのようにプログラムされているか確かめてみましょう。

leaf4# vsh_lc			
module-1# show plat	tform internal hal ep 13 all		
LEGEND:			
VrfName:	Vrf Name	т:	Туре
(Pl: Physical, Vl:	Virtual, Xr: Remote)		
EP IP:	Endpoint IP		
S Class:	S Class	Age Intvl:	Age
Interval			
S T:	Static Ep	S E:	
Secure EP			
L D:	Learn Disable	B N D:	Bind
Notify Disable			
E N D:	Epg Notify Disable	B E:	
Bounce Enable			
I D L:	IVxlan Dont Learn	SPI:	
Source Policy Incor	mplete		
DPI:	Dest Policy Incomplete	SPA:	
Source Policy Appl:	ied		
DPA:	Dest Policy Applied	DSS:	Dest
Shared Service			
IT:	Is Local	VUB:	Vnid
Use Bd			
SO:	SA Only	EP NH L3IfName:	EP
Next Hop L3 If Name	e		
NHT:	Next Hop Type (L2: L2 Entry L3: L3 Next Hop)	BD Name:	L2 NH
BD Name			
EP Mac:	EP Mac	L3 IfName:	L3 NH
If Name			
L2 IfName:	L2 If Name	FD Name:	L2
Entry FD Name			
ID:	L3 NH IP		
L3 EP Count: 12			
			=====
			=
	BEI	SDSDD V	EP-NH
N			
Vrf El	P S Age S S L N N B D	PPPPSIUS	L3
H BD EP	L3 L2 FD		
Name T II	P Class Intvl T E D D D E L	IIAASLBO	
IfName T I	Name Mac IfName Ifname Name	e IP	
=======================================			=====

==============					===	===	==	==	==	==:	===	==:	===	==	==	==	= =	==	==:	==	
common*rewall	Pl 10.6.112.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
common*rewall	Pl 10.6.114.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
common*rewall	Pl 10.6.114.129		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
common*efault	Pl 100.100.101.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	0								
Joey-T*ternal	Pl 192.168.1.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	0								
Joey-T*ternal	Xr 192.168.1.100		8	3013	128		0	0	0	1 (0 0	0	0	0	0	0	0	0	1 (0 -	-
L3 -	00:0c:0c:0c:0c:0c	Tunnel2	Tunr	nel2		-				0	.0.	0.0	0								
Joey-T*ernal2	Pl 192.168.3.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
Joey-T*ternal	Pl 192.168.20.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
Joey-T*ternal	Pl 192.168.20.2		8	300a	0		0	0	0	0	0 0	0	0	0	0	0	0	1	0	0 -	-
L2 BD-28	00:50:56:a5:fc:cc	-	Po3			FD-	30			-											
Joey-T*ternal	Pl 192.168.21.1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								
Joey-T*ternal	Pl 192.168.21.2		8	300c	0		0	0	0	0	0 0	0	0	0	0	0	0	1	0	0 -	-
L2 BD-7	00:50:56:a5:0c:11	-	Po4			FD-	8			-											
Joey-T*ternal	Pl 2001:0:0:100::1		1	L	0		1	0	0	0	0 0	1	1	0	0	0	0	1	0	0 -	-
L3 -	00:00:00:00:00:00	-	-			-				0	.0.	0.0	C								

ハードウェアはEPがTunnel 2上にあると認識しています。Tunnel 2の宛先は何ですか。

module-1# show system internal eltmc info interface tunnel2

trunk_id:

tunnel_mod:

tunnel_port:

ip_if_mode:

router_mac:00.00.00.00.00.00 unnumbered: 0

tep_ip: 0xc0a87843 p if mode: 0

IfInfo:				
interface:	Tunnel2	:::	ifindex:	402718722
iod:	66	:::	state:	up
Mod:	0	:::	Port:	0
Tunnel Index:	0	:::	Tunnel Dst ip:	0xc0a87843
Tunnel Encap:	ivxlan	:::	Tunnel VPC Peer:	0
Tunnel Dst ip str:	192.168.120.67	:::	Tunnel ept:	0x1
[SDK Info]:				
tunnl_name:				
vrf_id:	2	:::	if_index:	0x18010002
hwencapidx:	0	:::	encaptype:	1
mac_proxy:	0	:::	v4_proxy:	0
v6_proxy:	0	:::	ip_addr_type:	0
ipv4_address:	0xc0a87843			
[SDB INFO]:				
iod:	66			
<pre>pc_if_index:</pre>	0			
fab_if_index:	0			
sv_if:	0			
<pre>src_idx:</pre>	0			
int_vlan:	0			
encap_vlan:	0			
<pre>mod_port_status:</pre>	0x41620003			
v6_tbl_id:	0x8000002			
v4_tbl_id:	0x2			

0

0

0

0

sdk_vrf_id:	2			
mtu:	9366	:::	ipmtu_id:	0
с с <u>і</u> і .	0			

is_fex_fabric:

宛先は vPC の外部にあるため、その宛先 IP はリモート リーフの vPC の仮想 IP である必要があ ります。 リモート リーフを確認してみます。次をご覧ください。

leaf1# show system internal epm vpc

Local TEP IP	:	192.168.160.95
Peer TEP IP	:	192.168.160.93
vPC configured	:	Yes
VPC VIP	:	192.168.120.67
MCT link status	:	Up
Local vPC version bitmap	:	0x7
Peer vPC version bitmap	:	0x7
Negotiated vPC version	:	3
Peer advertisement received	:	Yes
Tunnel to vPC peer	:	Up

リモート vPC ペアから宛先 EP を学習しています。 ELAM が何を認識するかを確認し、パケッ トが正しく転送されていることを確かめてみましょう。

ELAM

ここで、EX ハードウェアのリモートの宛先には、パケット フローをトラブルシューティングす る際に非常に重要になる 2 つの ELAM の値があります。 ovector_idx は前の例と同様で、 encap_idx は次のようになります。

module-1(DBG-TAH-elam-insel6)# report | grep ovec sug_elam_out_sidebnd_no_spare_vec.ovector_idx: 0xB8 module-1(DBG-TAH-elam-insel6)# report | grep encap sug_lurw_vec.encap_l2_idx: 0x0 sug_lurw_vec.encap_pcid: 0x0 sug_lurw_vec.encap_idx: 0x6 sug_lurw_vec.encap_vld: 0x1

EX ハードウェアでは、パケットが転送される宛先ポートを有効にすることができます。 これま では、encap インデックスを確認し、宛先のインデックスが正しいトンネルであることを確認す るのみでした。 ここでは、どのポートが 8B にマップされているかを確認できます。

module-1(DBG-TAH-elam-insel6)# show platform internal hal 12 port gpd
Legend:
----IfId: Interface Id IfName:
I P: Is PC Mbr IfId:

Uc Pc Mbr Id Uc PC Cfg: UcPcCfg Idx Uc PC MbrId: AP: Asic Port As: Asic S1: Slice Sp: Slice Port Slice SrcId Ovector (slice | Ss: Ovec: srcid) L S: Local Slot Reprogram: T.3: Is L3 PifTable P: Xla Idx: Xlate Idx Rw PifTable Ovx Idx: OXlate Idx RP: IP: If Profile Table N L3: Num. of L3 Ifs RS: Rw SrcId Table NI L3: Num. of Infra L3 Ifs DP: DPort Table Vif Tid: Vif Tid RwVif Tid SP: SrcPortState Table RwV Tid: RSP: RwSrcPortstate Table Ing Lbl: Ingress Acl Label UC: UCPcCfq Egr Lbl: Egress Acl Label UM: UCPcMbr Reprogram: PROF ID: Lport Profile Id VS: VifStateTable HT: LportProfile Hw Install RV: Rw VifTable Num. of Sandboxes: 1 Sandbox_ID: 0, BMP: 0x0 Port Count: 8 _____ UC UC Reprogram | Rep | I PC PC L | RIRD R UUX | L Xla Ovx N NI Vif RwV Ing Egr | V R | PROF H Ifname PCfg MbrID As APSl Sp Ss Ovec S | PPPS PSp Sp CML | 3 Idx Idx L3 TfTd L3 Tid Tid Lbl Lbl | SV | ID I _____ _____ 1a004000 Eth1/5 1 0 1d 0 d 0 c 18 18 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 1a005000 Eth1/6 10 b 0 e 0 d 1a 1a 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - - 800 0 0 1 0 0 1a006000 Eth1/7 0 26 5 0 f 0 e 1c 1c 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 D-256 - 800 0 0 1 c 0 0 2f 7 0 10 0 f le le 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1a007000 Eth1/8 D-199 - 800 0 01 2e 0 1a01e000 Eth1/31 10 2d 0371 e 1c9c 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - 0 0 0 1 0 0 1a01f000 Eth1/32 1 0 3d 0 38 1 f 1e 9e 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - 0 0 0 1 0 0 1a030000 Eth1/49 0 2 1 0 49 1 20 38 b8 1 0000000000164 2 2 D-24d - 400 0 001 0 1a031000 Eth1/50 0 3 3 0 29 1 0 0 80 1 0 0 0 0 0 0 0 0 0 1 5 3 2 2 0 0 D-350 - 400 0 1

スイッチはインターフェイスEth1/49のスパインに転送すべきと考えています。しかし、encapが 正しいことを確認するにはどうすればよいのですか。

まず、トンネルに関するハードウェア情報を確認する必要があります。 次の HAL コマンドを実 行して、これを行えます。

module-1(DBG-TAH-elam-insel6)# show platform internal hal tunnel rtep pi Non-Sandbox Mode LEGEND:

Tun Ifid: Tunnel Ifid IfName: Tunnel If Name Logical Id ET: Lid: Encap Type V: Vxlan I: IVxlan N: NVGRE VrfId: Vrf Id Vrf Name: Vrf Name TP: Tunnel's IP Hw Enc: Hw Encap Idx TVP: Is VPC Peer Is Local Proxy for v4 TT.: P4: Proxy for V6 P6: PM: Proxy for Mac Is Ingress Only IC: Is Copy Service TT: Copy Service Outer Bd Use DF C OBd: UD: Next Base Type E: ECMP N: Next-Hop NB Id: Next Base Id NBT: NH cnt: Next Hop Count VrfId: Vrf Id Vrf Name: Vrf Name TP: TP Address L3 IfId: L3 IfId Mac: Mac L3IfName: L3 If Name L2 IfId: L2 IfId L2IfName: L2 If Name Num. of Sandboxes: 1 Sandbox_ID: 0, BMP: 0x0 Remote Tep Count: 15 _____ _____ ======= ΝΝ I HW VIPPPIIC UBB E Vrf NH Vrf T.3 T.3 T.2 T.2 IP T Lid VrfId Name Enc PL46MICOBdDTId TfTd Ifname Cnt | VrfId Name IP IfId IfName IfId Mac IfName _____ _____ ======= **18010002 Tunnel2** I 3005 2 overlay-1 192.168.120.670 0 0 0 0 0 0 0 1 0 E 2 0d:0d:0d:0d:0d:00 1a030001 Eth1/49.1 1a030000 Eth1/4 overlay-1 0.0.0.0 2 2 9 overlay-1 0.0.0.0 0d:0d:0d:0d:0d:00 1a031002 Eth1/50.2 1a031000 Eth1/5 2 Ω この出力によって、ここで必要としている以下の値が提供されます。 IfId - トンネルに割り当てられたインターフェイス ID。 IP: 宛先のIP。 これは ELTMC と一致している必要があります。 L3 IfId - スイッチが適切な宛先に転送するために使用できるレイヤ3インターフェイス。 IfId を把握したら、ELAM で取得した encap がトンネルの宛先と一致していることを確認できま す。

module-1(DBG-TAH-elam-insel9)# show platform internal hal tunnel rtep apd
Non-Sandbox Mode
LEGEND:
-----ifId: Interface Id IP: I
HwVrfId: Hardware Vrf Id SrcTepIdx: S

BDXlate: RwEncapIdx: Num: L2 Index:	Egress BDX] Rw Encap Ir Number of P L2 Index	late Idex Iops				DstInfoIdx: ECMPIdx: ECMPMbrIdx: RwDmacIdx:	Destin ECMP I ECMP m Rw Dma	ation info index member Inde x Index	index x
Num. of Sand	boxes: 1								
Sandbox_ID: Remote Tep	0, BMP: 0x0 Count: 15								=====
ifId IP L2Index RwD	macIdx	HwVrfId	BDXlate	SrcTepIdx	DstInfoId>	K RwEncapIdx	ECMPIdx	ECMPMbrId	x Num
18010002 192 1a030000 0	.168.120.67 <	2 - RwEncap	1 Jdx is (3a9a 5! Same a	3005 s the "enca	6 ap_idx" in th	0 ne ELAM R	0 Peport.	2

1a031000 1

このトンネルには、値 6 の RwEncapIdx(Encap インデックスの書き換え)があります。これは、ELAM に表示されている内容です。

1 EP ---> L3アウト - ルーテッドフロー

トポロジ



EP1 N5K -OSPF 0050.56a5.50ab 192.168.20.10/24 100.100.100/32

この例では、OSPFを実行しているN5KのループバックにICMPを送信するEP1からのパケットの パケットフローを追跡します。N5Kは、同じEXスイッチペア上のL3Out経由で接続されます。

このドキュメントの最初で、ローカル EP プログラミングについて確認しているため、ここでは

ハードウェアで EP が正しく学習されていることを前提とし、ルートの検証に話を進めましょう。

最初に、OSPFの状態とルーティングテーブルを確認してみましょう。

leaf6# show ip ospf neighbors vrf jr:sb OSPF Process ID default VRF jr:sb Total number of neighbors: 2 Neighbor ID Pri State Up Time Address Interface 27.27.27.1 1 FULL/BDR 00:22:39 10.10.27.1 Vlan28 <---- Leaf5 27.27.27.3 1 FULL/DROTHER 00:22:37 10.10.27.3 Vlan28 <---- N5K</pre>

leaf6# show ip route vrf jr:sb 100.100.100
IP Route Table for VRF "jr:sb"
'*' denotes best ucast next-hop
'**' denotes best mcast next-hop
'[x/y]' denotes [preference/metric]
'%<string>' in via output denotes VRF <string>

100.100.100.100/32, ubest/mbest: 1/0

*via 10.10.27.3, vlan28, [110/5], 00:16:58, ospf-default, intra

ルーティングテーブルに10.10.27.3の5Kのネクストホップが示されていることがわかっています 。良いスタートですが、どのハードウェアが搭載されているかを確認するにはどうしたらいいで すか。

まず、ハードウェアの隣接関係テーブルをチェックし、ARP が 10.10.27.3 に解決されていること、また正しいインターフェイスにプログラムされていることを確認します。

leaf6# vsh_lc module-1# show forwarding adjacency

IPv4 adjacency information, adjacency count 20

MAC アドレスは	ま、次のように	5K 上のアドレス	くと一致しています。
10.10.27.3	8c60.4f02.88fc	Vlan28	port-channel5
10.10.27.1	0022.bdf8.19ff	Vlan28	Tunnel3
next-hop	rewrite info	interface	phy i/f

ACI-5548-B# show interface vlan 3117

Vlan3117 is up, line protocol is up Hardware is EtherSVI, address is 8c60.4f02.88fc Internet Address is 10.10.27.3/29 MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec

EXプラットフォームには、VRFに割り当てられた「hw_vrf_idx」があります。このインデックス は、ハードウェアプログラミングを確認するときに参照されます。 インデックスを検索してみま しょう。

module-1# show system	internal eltmc	info vrf	jr:sb	
VRF-TABLE: jr:sb				
vrf_type:	tenant	:::	context_id:	6
overlay_index:	0	:::	vnid:	2129921
scope:	5	:::	sclass:	16386
v4_table_id:	0x5	:::	v6_table_id:	0x80000005
intf_count:	5	:::	intrn_vlan_id:	0

VRF Intf:	Vlan11	:::	<pre>src_plcy_incomp:</pre>	0	
vnid_hex:	0x208001	:::	ingress_policy:	0x1	
vrf_intf_list:	Vlan28,Vlan16,Vla	an9,	Vlan11,loopback2,		
hw_vrf_idx:	4612	:::	nb_egr_outer_bd:	0	
sb_egr_outer_bd:	0				
vrf_bd_list:	28,16,11,9,				
sb_egr_outer_bd:	0	:::	sdk_vrf_id:	5	
[SDK Info]:					
vrf_name:	jr:sb				
vrf_id:	5	:::	hw_vrf_idx:	4612	
vrf_vnid:	2129921	:::	is_infra:	0	
tornbinfrahwbd:	0	:::	torsbinfrahwbd:	0	
ingressBdAclLabel:	0	:::	ingBdAclLblMask:	0	
egressBdAclLabel:	0	:::	egrBdAclLblMask:	0	
sg_label:	5	:::	sclass:	16386	
sp_incomplete:	1	:::	sclassprio:	3	
[SDB INFO]:					
v4 table					
vrf type:	1				
vrf id:	5				
vnid:	2129921				
internal infra vlan:	0				
external router mac:	00:22:bd:f8:19:ff				
v6 table					
vrf type:	1				
vrf id:	5				
vnid:	2129921				
internal infra vlan:	0				
external router mac:	00:22:bd:f8:19:ff				
::::					

隣接関係を検出すると、HAL はルートをプログラムします。 次のコマンドを使用してこれを確 認できます。

module-1# show platform internal hal 13 routes | head _____ _____ LEGEND: _____ _____ PID: Physical ID NB-ID:Next-Base ID LID: Logical ID RID: Route ID HIT IDX: Next-Hop HitIndexCLP : Class PriorityTBI: Trie Base Index|SC : Sup-CopySSR: Src Sup-RedirectDSR: Dst Sup-Redirect TDD :TTL Disable TBI: Trie Base Index NB: NextBaseType SDC : Src Direct Connect TRO: Trie Offset NB: NextBaseTypeSDC : Src Direct ConnectTRO: Trie Offset|SPI: Src Policy IncDPI: Dst Policy IncDR : Default RouteLE :Learn Enable ILL : Is Link Local ISS: Is Shared Services [E:Ecmp/A:Adj] HR : Host Routes EP RT : Route Type FWD: Forwarding :Ext Prefixes CLSS: Class Id 1 DLR: Default Lpm Route RDEL: Route in Deletion BNE: Bind Notify Enable SNE: Sclass Notify Enable BE : Bounce Enable IDL : Ivxlan DoNotLearn DL : Dest Local SA : Src Only AI : Age Interval SF : Static Flag SH : Src Hit DH: Dest Hit module-1# show platform internal hal 13 routes _____ _____ LEGEND:

_____ PID: Physical ID NB-ID:Next-Base ID LID: Logical ID RID: Route ID HIT IDX: Next-Hop HitIndexCLP : Class PriorityTBI: Trie Base Index|SC : Sup-CopySSR: Src Sup-RedirectDSR: Dst Sup-Redirect TDD :TTL Disable NB: NextBaseType SDC : Src Direct Connect TRO: Trie Offset | DPI: Dst Policy Inc DR : Default Route LE :Learn Enable SPI: Src Policy Inc ILL : Is Link Local ISS: Is Shared Services [E:Ecmp/A:Adj] RT : Route Type FWD: Forwarding HR : Host Routes EP :Ext Prefixes RDEL: Route in Deletion DLR: Default Lpm Route CLSS: Class Id BNE: Bind Notify Enable SNE: Sclass Notify Enable BE : Bounce Enable IDL : Ivxlan DoNotLearn DL : Dest Local SA : Src Only AI : Age Interval SF : Static Flag SH : Src Hit DH: Dest Hit _____ | LID |<------ Trie ----->|<Dleft Trie>| | RT| RID | LID | Type| PID | FPID/| HIT VRF Prefix/Len N NB-ID NB Hw | PID | FPID/ | TBI | TRO | Ifindex | CLSS | CLP | AI | SH | DH | Flags --|-----|----|---|---| TID | IDX | - - -| PID | FPID/| HIT N NB-ID NB Hw | | | TID | IDX |B| Idx | | . . | |<-----| - TCAM ----->| | | | PID | TCAM | HIT N NB-ID NB Hw | ID | IDX |B| | **Idx** | _____ _____ Sandbox_ID: 0 Asic Bitmap: 0x0 _____ _____

module-1# show platform internal hal 13 routes | egrep 100.100.100.100 | 4612| 100.100.100/32| UC| e4| 4a04| TRIE| 10| 5/ 0| 6010|A| 7567| 802e| 186a| 1/ 2| 10| 0| 0| f| 3| 0| 0| 0|spi,dpi この出力は、ネクスト ホップのルートに関する情報を示します。 4612は、jr:sb VRFの hw_vrf_idxです。ネクストホップを確認するために、TCAMの「NB Hw Idx」が次の表に対して使 用されます。

Non-Sandbox Mode LEGEND: _____ NHOP ID : Nhop Identifier (Hex) CONS : H/W S/W info Consistency : Nexthop Type ACTN : Nexthop Action TYPE L3 INTF : L3 interface index (Hex) : L3 Vrf of the Nhop Vrf L2 INTF : L2 interface index (Hex) BDID Or RwVRF : Bridge Domain Id Or Rewrite Vrfid (Hex) : ACI Infra valid PVRF INFR : Preserve VRF : Learn Enabled : VRF Rewrite LRN VRFR FPID : FP of this nexthop : Physical ID PID TLID : Tile Id within FP HIT IDX : Location of this Nhop (Hex)

module-1# show platform internal hal 13 nexthops

Mac Entry: : Type INTF TYP : Interface related Info (Hex) LRN : Learn Info DL : Destination Local : Unused VNB : Vnid use BD MLD DFL : Default Entry VLD : MacKey Valid : FID Type FV : FID Valid FT : FID value (Hex) Mac FID : L2 MAC Address L2 Ifabric Info: CLSS : Source Class : Source Class Priority CLP : Bind Notification Enabled EPG : EndPoint Group BNE : Source Address Notification Enabled CNE : Source class Notification SNE Enabled : iVxlan DL SPT : Source Policy Incomplete DL DPT : Dest Policy Incomplete IP Address : IP address Sandbox_ID: 0 Asic Bitmap: 0x0 Summary info for 31 L3 Nexthop objects BDID I P V T |-----Mac Entry-СТА -----L2 Ifabric Info----| NHOP OYC L3 L2 Or NVLR LHIT|T L MVDV|-----------Mac Key-----| С SD вsс ID N P T INTF INTF RWVRFFRRF FPIIDX YINTF RDLNFL FFID L NNNDPP (Hex) S E N Vrf (H) (H) (H) R F N R PID ID D (H) P (H) N L D B L D T V (H) CLSS P EPG E E E L I I IP Address Mac

module-1# show platform internal hal 13 nexthops | grep 802e 7567 NIF 5 901001c 1600004 1c 0 0 0 0 2e 9 0 802e 0 22 0 0 0 0 0 1 1 1 1214 8c:60:4f:02:88:fc 0 0 2c0d 0 0 0 0 0 10.10.27.3 ここでは、「NB Hw Idx」を取り、「HIT IDX」にマッピングします。これは、ネクストホップ MAC/IPに対応するエントリを示します。これは、Generation 1 ACIリーフスイッチの Broadcomの「I3 definp show」および「I3 egress show」をで見同等です。

ご覧のとおり、テーブルには正しい情報が示されています。

L2 INTF: 0x16000004 —>ポートチャネル5のifIndex

HIT IDX: hal I3 ルートの Nb Hw Idx からのインデックス

MAC: 8c:60:4f:02:88:fc —>ネクストホップSVIの5K上のMAC

EPG: L3 EPG の SCLASS

IP アドレス: 10.10.27.3 —> 5K上のSVIのネクストホップIP

ELAM

leaf6# pwd
/var/sysmgr/tmp_logs

 leaf6# cat elam_report.txt | grep adj
sug_lurw_vec.dst_addr.adj: 0x8C604F0288FC
sug_lurw_vec.dst_addr.adj.padfield: 0x04F0288FC
sug_lurw_vec.dst_addr.adj.idx: 0x2318
sug_lurw_vec.adj_vld: 0x0

1 EP —>リモートEPまたはSVI – スパイン検証

トポロジ



EP1 EPG1 10.100.17.11

ロジック

この例では、EP1からRemote BD Switched Virtual Interface(SVI)宛てのパケットのパケットフローを追跡します。この例の目的は、スパイン転送を確認して、パケットが正しいリーフに送信されることを確認することです。 パケットが入力リーフのスパイン プロキシに送信されたことを前提とします。

パケットはルックアップのためにスパイン プロキシに送信されるため、スパインでまず宛先 IP の Council of Oracles Protocol(COOP)を確認します。

calo1-spine1# show coop internal info ip-db | grep -A 10 192.168.20.1 IP address : 192.168.20.1 Vrf : 2129921 Flags : 0 EP vrf vnid : 2129921 EP IP : 192.168.20.1 Publisher Id : 10.0.224.88 Record timestamp : 11 04 2016 16:41:16 422062712 Publish timestamp : 11 04 2016 16:41:16 424633605 Seq No: 0 Remote publish timestamp: 01 01 1970 00:00:00 0 URIB Tunnel Info Num tunnels : 1 Tunnel address : **10.0.224.88** <---- REMOTE LEAF Tunnel ref count : 1 どのリーフがその TEP アドレスを持つか確かめてみましょう。

spinel# acidiag fnvread | grep 10.0.224.88 105 1 calol-leaf5 FD020160TPS 10.0.224.88/32 leaf active 0 パケットは Module 2、Port 6 のスパインで受信されることが分かっているため、Module 2 に接 続して、ポート レイアウトを確認します。

spine1# vsh Cisco iNX-OS Debug Shell This shell should only be used for internal commands and exists for legacy reasons. User should use ibash infrastructure as this will be deprecated. calo1-spine1# attach module 2 Attaching to module 2 ... To exit type 'exit', to abort type '\$.' No directory, logging in with HOME=/ Bad terminal type: "xterm-256color". Will assume vt100. Cisco iNX-OS Debug Shell This shell should only be used for internal commands and exists for legacy reasons. User should use ibash infrastructure as this will be deprecated. Loading parse tree (LC). Please be patient... module-2#

```
module-2# show platform internal hal 12 port gpd
Legend:
```

IfId:	Interface Id	IfName:	Interface Name
I P:	Is PC Mbr	IfId:	Interface Id
Uc PC Cfg	: UcPcCfg Idx	Uc PC MbrId:	Uc Pc Mbr Id
As:	Asic	AP:	Asic Port
sl:	Slice	Sp:	Slice Port
Ss:	Slice SrcId	Ovec:	Ovector (slice
srcid)			
L S:	Local Slot	Reprogram:	
L3:	Is L3		
P:	PifTable	Xla Idx:	Xlate Idx
RP:	Rw PifTable	Ovx Idx:	OXlate Idx
IP:	If Profile Table	N L3:	Num. of L3 Ifs

RS: Rw SrcId Table NT L3: Num. of Infra L3 Ifs DPort Table Vif Tid Vif Tid: DP: RwV Tid: SP: SrcPortState Table RwVif Tid RSP: RwSrcPortstate Table Ingress Acl Label Ing Lbl: UC: UCPcCfg Egr Lbl: Egress Acl Label UM: UCPcMbr Reprogram: PROF ID: Lport Profile Id VifStateTable LportProfile Hw VS: HT: Install RV: Rw VifTable Num. of Sandboxes: 1 Sandbox_ID: 0, BMP: 0x0 Port Count: 7 _____ Reprogram UC UC Rep I PC Pc L | RIRD R UUX | L Xla Ovx N Ing Egr | V R | PROF H NI Vif RwV P Cfg MbrID As AP Sl Sp Ss Ovec S | P P P S P Sp Sp C M L | 3 Idx Idx L3 IfId Ifname Tid Lbl Lbl | SV | ID L3 Tid I _____ -----1f5 D-2d4 D-3e1 0 0 0 0 1 0 0 9a 1c 0 11 0 10 20 20 1 0 0 0 0 0 0 0 0 0 1b b 1 1 1a080000 Eth2/1 D-f3 D-61 100 0 1 0 0 0 1a081000 Eth2/2 0 9b 22 0 d 0 c 18 18 1 0 0 0 0 0 0 0 0 0 0 1 c С 1 1 0 0 D-lee D-30b 100 0 0 1 0 9e 1e 0 3d 1 14 28 a8 1 1a084000 Eth2/5 0 0 0 0 0 0 0 0 0 1 1 1 1 1 D-19a D-2ee 100 0 0 0 1 0 0 9f 24 **0** 39 **1** 10 20 a0 1a085000 Eth2/6 1 0 0 0 0 0 0 0 0 0 1 e е 1 1 D-87 D-184 100 0 0 0 1 0 0 a0 26 0 35 1 c 18 98 0 0 0 0 0 0 0 0 0 0 1a086000 Eth2/7 1 1d d 11 D-1d0 D-357 100 0 0 0 1 0 1a088000 Eth2/9 0 a2 20 0 0 D-3ea D-1a9 100 0 1 Ο

Ethernet 2/6は、Leaf 6に接続するインターフェイスはASIC 0 SLICE 1上にあります

これで、どの ASIC で ELAM を実行するかが分かりました。 ASIC 0 です。

module-2# debug platform internal tah elam asic 0 module-2(DBG-TAH-elam)# trigger reset module-2(DBG-TAH-elam)# trigger init in-select 13 out-select 0 module-2(DBG-TAH-elam-insel13)# set inner ipv4 src_ip 10.100.17.11 dst_ip 192.168.20.1 module-2(DBG-TAH-elam-insel13)# start stat module-2(DBG-TAH-elam-insel13)# stat ELAM STATUS _____ Asic 0 Slice 0 Status Armed Asic 0 Slice 1 Status Armed module-2(DBG-TAH-elam-insel13)# stat ELAM STATUS _____ Asic 0 Slice 0 Status Triggered <---- Packet triggered from FM Asic 0 Slice 1 Status Triggered <---- Packet triggered from Front Panel ELAM を確認すると、ovector インデックスが分かります。

Front Panel ELAM drove sug_elam_out_sidebnd_no_spare_vec.ovector_idx: 0xB8 さて、どのように 0xb8 をポートにマッピングしたらよいでしょうか。 パケットはルックアップ のためにファブリック モジュール(FM)に送信されるため、内部ポート マッピングを調べて宛 先 FM を見つけます。

module-2# Num. of Sa Legend:	show platform internal hal 1 ndboxes: 1	l2 internal-p	oort pi
IfId:	Interface Id	IfName:	Interface Name
As:	Asic	AP:	Asic Port
sl:	Slice	SP:	Slice Port
Ss:	Slice SrcId	Ovec:	Ovector
UcPcCfgId:	Uc Pc CfgId	Lb Mbrid:	LB MbrId
Sandbox_ID	: 0, BMP: 0x0		
Internal	Port Count: 32		
			UcPc Lb
	Ca.		

IfId	IfName	As	AP	Sl	SP	Ss	Ovec	CfgId	MbrId
========		===	===	===	===	===	=====		
7d	-	0	21	0	20	38	38	0	4
7e	-	0	29	1	0	0	80	0	8
7f	-	1	21	0	20	38	38	0	С
80	-	1	29	1	0	0	80	0	10
81	-	2	21	0	20	38	38	0	14
82	-	2	29	1	0	0	80	0	18
83	-	3	21	0	20	38	38	0	1c
84	-	3	29	1	0	0	80	0	20
95	-	0	19	0	18	30	30	0	3
96	-	04	19 1	2	20 3	88 I	o8 () 7	
97	-	1	19	0	18	30	30	0	b
98	-	1	49	1	20	38	b8	0	f
99	-	2	19	0	18	30	30	0	13
9a	-	2	49	1	20	38	b8	0	17
9b	-	3	19	0	18	30	30	0	1b
9c	-	3	49	1	20	38	b8	0	1f
ad	-	0	25	0	24	40	40	0	1
ae	-	0	41	1	18	30	b0	0	6
af	-	1	25	0	24	40	40	0	9
b0	-	1	41	1	18	30	b0	0	е
b1	-	2	25	0	24	40	40	0	11
b2	-	2	41	1	18	30	b0	0	16
b3	-	3	25	0	24	40	40	0	19
b4	-	3	41	1	18	30	b0	0	le
dd	-	0	15	0	14	28	28	0	2
de	-	0	4d	1	24	40	с0	0	5
df	-	1	15	0	14	28	28	0	a
e0	-	1	4d	1	24	40	с0	0	d
el	-	2	15	0	14	28	28	0	12
e2	-	2	4d	1	24	40	с0	0	15
e3	-	3	15	0	14	28	28	0	1a
e4	-	3	4d	1	24	40	с0	0	1d
ASIC0 /	Ovec B8を使用して、Mbrld	0x7	を	取得	寻し	ま	す。		

この Mbrld は、FM 上のインターフェイスにマッピングされる USD のインターフェイスです。 このMbrldは16進数であり、10進数に変換する必要があることに注意してください。 module-2# show platform internal usd port info | grep -A 3 "Int 7"(if the interface has multiple digits, will be "Int##" with no space)

Port 73.0 (Int 7) : Admin UP Link UP Remote slot22.asic0
slice:1 slice port:32 lcl srcid:56 gbl srcid:184
asic mrl:0xd07c010, mac mrl:0x12c84010, mac:16, chan:0
speed 106G serdes: 0x328 0x329 0x32a 0x32b

「スロット」は 0 から始まりますが、FM の番号付けは 1 から始まります。そのため、ここでリ ストされている番号に 1 を追加する必要があります。 つまり、パケットは FM 23 に送信されま す。

シンセティック IP

Alpine と同様、COOP ルックアップ用のハッシュを決定するための、外部用 IP アドレスとして 使用されるシンセティック IP があります。 これを検出するには、次のコマンドを実行して、内 部用 DST IP を grep する必要があります。

module-2(DBG-TAH-elam-insel7)# show forwarding route synthetic vrf all | grep 192.168.20.1 SYNTH-88 1.203.211.185/32 0x208001 192.168.20.1 これは、1.203.211.185が合成IPであることを示しています。これをもとに、FM elamの「Outer DST IP」を設定することもできます。 次のように、FM 上でトリガーする必要があります。

ファブリック モジュール ELAM

```
module-23(DBG-TAH-elam-insel7)# trigger reset
module-23(DBG-TAH-elam)# trigger init in-select 13 out-select 0
module-23(DBG-TAH-elam-insel13)# set outer ipv4 dst_ip 1.203.211.185 <---- DST IP IS THE
SYNTHETIC IP
module-23(DBG-TAH-elam-insel13)# set inner ipv4 src_ip 10.100.17.11 dst_ip 192.168.20.1
module-23(DBG-TAH-elam-insel13)# start
stat
module-23(DBG-TAH-elam-insel13)# stat
ELAM STATUS
============
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Armed
Asic 0 Slice 2 Status Armed
Asic 0 Slice 3 Status Armed
Asic 0 Slice 4 Status Armed
Asic 0 Slice 5 Status Armed
module-23(DBG-TAH-elam-insel13)# stat
ELAM STATUS
_____
Asic 0 Slice 0 Status Armed
Asic 0 Slice 1 Status Armed
Asic 0 Slice 2 Status Triggered
                               <---- Triggered on SLICE 2</pre>
Asic 0 Slice 3 Status Armed
Asic 0 Slice 4 Status Armed
Asic 0 Slice 5 Status Armed
このように、レポートがすべてダンプされますが、トリガーしたこのパケットの ovector_idx を
確認します。
```

lac_elam_out_sidebnd_no_spare_vec.ovector_idx:**0x20** < – 次のコマンドで使用されるOvector Index

どのインターフェイスにその ovector があるか、どのように見つけ出すことができますか。 FM で次を実行します。

**バグ<u>CSCvf42796</u> 、すべてのFMコマンドを「| no-more」と追加します。 それ以外の場合、最 終的な出力に特定のテーブルエントリが表示されないことがあります。

module-23(DBG-TAH-elam-insel13)# show platform internal hal 12 port gpd | no-more Legend: ____ IfId: Interface Id IfName: Interface Name Is PC Mbr IfId: I P: Interface Id Uc PC Cfg: Uc PC MbrId: Uc Pc Mbr Id UcPcCfg Idx As: Asic AP: Asic Port Slice Sp: Slice Port Sl: Slice SrcId Ovec: Ovector (slice | Ss: srcid) L S: Local Slot Reprogram: L3: Is L3 P: PifTable Xla Idx: Xlate Idx RP: Rw PifTable Ovx Idx: OXlate Idx IP: If Profile Table N L3: Num. of L3 Ifs RS: Rw SrcId Table NI L3: Num. of Infra L3 Ifs DP: DPort Table Vif Tid: Vif Tid SP: SrcPortState Table RwV Tid: RwVif Tid RSP: RwSrcPortstate Table Ing Lbl: Ingress Acl Label UC: UCPcCfg Egr Lbl: Egress Acl Label UM: UCPcMbr Reprogram: PROF ID: Lport Profile Id VS: HI: VifStateTable LportProfile Hw Install RV: Rw VifTable Num. of Sandboxes: 1 Sandbox_ID: 1, BMP: 0x1 Port Count: 8 UC UC Reprogram | Rep | I PC Pc L | RIRD R UUX | L Xla Ovx N NI Vif RwV Ing Egr | V R | PROF H TfTd P Cfg MbrID As AP Sl Sp Ss Ovec S | P P P S P Sp Sp C M L | 3 Idx Idx L3 Ifname Lbl Lbl | SV | ID I L3 Tid Tid ______ ae fc0-lc1:0-0 1 0 3 0 fc0-lc1:0-1 1 0 4 0 3d 2 c 18 98 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 af 0 0 0 0 0 0 _ 0 d 0 c 18 18 0 0 0 0 0 0 0 0 0 b0 fc0-lc1:1-0 1 0 13 1 0 0 0 0 0 0 0 0 0 0 0 b1 fc0-lc1:1-1 1 0 14 0 39 2 8 10 90 1 0 _ fc0-lc1:2-0 1 0 23 0 5d 3 14 28 e8 0 0 0 0 0 0 0 0 0 0 0 0 0 0 b2 1 0 0 0 0 0 0 0 fc0-lc1:2-1 1 0 24 0 21 1 8 10 50 1 0 0 0 0 0 0 0 0 0 0 b3 0 0 0 0 0 0 0 0 0 0 0

fc0-lc1:3-0 1 0 b4 0 0 0 0 0 0 このovectorは、ASIC 0/SLICE 0上のLC1(スロット2のラインカードは0ベースであるため)にマ ッピングされます。ELAMがLC上で最初に実行していることがわかるように、次のスライスでト リガーされます。 module-2# debug platform internal tah elam asic 0 module-2(DBG-TAH-elam)# trigger reset module-2(DBG-TAH-elam)# trigger init in-select 13 out-select 0 module-2(DBG-TAH-elam-insel13)# set inner ipv4 src_ip 10.100.17.11 dst_ip 192.168.20.1 module-2(DBG-TAH-elam-insel13)# start stat

module-2(DBG-TAH-elam-insel13)# stat

ELAM STATUS

Asic 0 Slice 0 Status Armed Asic 0 Slice 1 Status Armed

module-2(DBG-TAH-elam-insel13)# stat
ELAM STATUS
=========

Asic 0 Slice 0 Status Triggered <---- Packet triggered from FM

Asic 0 Slice 1 Status Triggered <---- Packet triggered from Front Panel

この ELAM の ovector は sug_elam_out_sidebnd_no_spare_vec.ovector_idx:**0x98 です。これは、** 「hal I2 port gpd」から分かりますが、LC 上の適切なインターフェイスにマッピングされます。

UC UC Reprogram | Rep | I PC Pc L | RIRD R UUX | L Xla Ovx N NI Vif Ing Egr | V R | PROF H RwV P Cfg MbrID As AP Sl Sp Ss Ovec S | P P P S P Sp Sp C M L | 3 Idx Idx L3 IfId Ifname Tid Lbl Lbl | SV | ID L3 Tid I _____ 1f5 1 D-2d4 D-3e1 0 0 0 0 0 0 11 0 10 20 20 1 1 1a080000 Eth2/1 0 9a 1c 1 0 0 0 0 0 0 0 0 0 1 b b D-f3 D-61 100 0 0 0 1 0 1a081000 Eth2/2 09b 22 0 d 0 c 18 18 0 0 0 0 0 0 0 0 0 1 1 1 lc c D-lee D-30b 100 0 0 0 1 0 0 9e 1e 0 3d 1 14 28 a8 0 0 0 0 0 0 0 0 0 1a084000 Eth2/5 1 1 1 1 1 1 D-19a D-2ee 100 0 0 0 1 0 0 39 1 10 20 a0 1a085000 Eth2/6 0 9f 24 0 0 0 0 0 0 0 0 0 1 1 e e 1 1 D-87 D-184 100 0 0 0 1 0 1a086000 **Eth2/7** 0 a0 26 0 35 1 c 18 **98** 1 0 0 0 0 0 0 0 0 0 1 d d 1 1 D-0 0 1 0 1d0 D-357 100 0 1a088000 Eth2/9 D-3ea D-1a9 100 0 0 0 1 0 Ethernet 2/7は、リーフ5に接続するインターフェイスです。

追加シナリオ:「hal internal-port pi」出力にない Ovector の取得

トポロジ



ロジック

「show platform internal hal l2 internal-port pi」テーブルに Ovector がないパケットを捕捉するい くつかのシナリオがあります。 以下のシナリオでは、実際には FM から戻ってくるパケットを捕 捉するため、別のテーブルを確認して、パケットがどの前面パネル ポートを選択しているかを確 かめる必要があります。

上記のトポロジは、中継トラフィックを学習する(プロキシ ルーティングなし)、まったく異な る環境のものです。 モジュールは N9K-X9732C-EX です。

@module-1(DBG-elam-insel13)# report | grep ovector sug_elam_out_sidebnd_no_spare_vec.ovector_idx: 0xA0 <<<<<<<< > now we look for this in the "hal internal-port pi" command @module-1# show platform internal hal 12 internal-port pi No sandboxes exist Num. of Sandboxes: 1 Legend: _____ Interface Name IfId: Interface Id IfName: As: Asic AP: Asic Port sl: Slice SP: Slice Port Slice SrcId Ovec: Ss: Ovector UcPcCfgId: Uc Pc CfgId Lb Mbrid: LB MbrId Sandbox_ID: 0, BMP: 0x0 Internal Port Count: 24 _____ UcPc Lb As AP Sl SP Ss Ovec CfgId MbrId IfId IfName _____ 0 21 0 20 38 38 0 7d _ 4 7e 0 29 1 0 0 80 0 8 1 21 0 20 38 38 0 7f _ С 1 29 1 0 0 80 0 80 _ 10 81 2 21 0 20 38 38 0 14 82 2 29 1 0 0 80 0 18 _ 3 21 0 20 38 38 0 83 _ 1c 84 3 29 1 0 0 80 0 _ 20 ad 0 25 0 24 40 40 0 1 0 41 1 18 30 b0 0 _ 6 ae 1 25 0 24 40 40 0 af _ 9 1 41 1 18 30 b0 b0 0 _ е b1 2 25 0 24 40 40 0 11 _ 2 41 1 18 30 b0 0 b2 _ 16 b3 _ 3 25 0 24 40 40 0 19 3 41 1 18 30 b0 0 b4 1e 0 15 0 14 28 28 0 dd 2 0 4d 1 24 40 c0 0 de _ 5 1 15 0 14 28 28 df 0 а e0 1 4d 1 24 40 c0 0 d 2 15 0 14 28 28 0 12 e1 _ e2 _ 2 4d 1 24 40 c0 0 15 3 15 0 14 28 28 0 e3 1a 3 4d 1 24 40 c0 0 ld <<<<<< we cant find an e4 entry that matches 0xA0 @module-1# show platform internal hal 12 port gpd Legend: _____ <snip> Sandbox_ID: 0, BMP: 0x0 Port Count: 6 _____ UC UC Reprogram | Rep |

Asic 3 Slice 1 Status Triggered

I	PC Pc					I	- -	R	ΙI	R D		R	υt	JZ	K	LΣ	Kla	0vx 1	N NI
Vif RwV Ing Egr V R PROF H																			
IfId Ifname	P Cfg Mb	orID As	AP	Sl Sp	Ss	Ovec	S	ΡI	? P	S	P S	o Sp	С	М	L	3	Idx	Idx	L3
L3 Tid Tid Lbl	Lbl S	V I	D	I															
		======	====																
1f5 SpInBndMgmt	0 9de 1a	. 0	0	0 0	0	0	0	0 0	0 0	0	0 0	0	0	0	0	0	0	0	0
0 D-2d4 D-3e1 0	0 0	0 1		0															
1a000000 Eth1/1	0 1b 1c	. 0	11	0 10	20	20	1	0 0	0 0	0	0 0	0	0	0	0	1	1	1	1
1 D-13b D-33b 500	0 1	0 3		0															
1a01c000 Eth1/29	0 37 le	: 3	3d	1 14	28	a8	1	0 0	0 0	0	0 0	0	0	0	0	1	8	8	1
1 D-3f2 D-7a 100	0 0	0 2		0															
1a01d000 Eth1/30	0 38 20	3	39	1 10	20	a0	1	0 0	0 0	0	0 0	0	0	0	0	1	5	5	1
1 D-36e D-362 100	0 0	0 2		0															
1a01e000 Eth1/31	0 39 22	3	35	1 c	18	98	1	0 0	0 0	0	0 0	0	0	0	0	1	9	9	1
1 D-273 D-8 100	0 0	0 2		0															
1a01f000 Eth1/32	03a 24	3	31	1 8	10	90	1	0 0	0 0	0	0 0	0	0	0	0	1	а	a	1
1 D-154 D-5d 100	0 0	0 2		0															

1/30は、リーフ102に接続する物理インターフェイスで、トポロジによって確認されます。ASIC 3、スライス1