

Nexus 7000 F2/F2e:了解并缓解MAC表完整问题

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[背景信息](#)

[缓解步骤](#)

[选项1.修剪VLAN](#)

[选项2. L3分离](#)

[选项3. Fabricpath等备用设计架构](#)

[选项4.使用M2/F3卡等高容量线卡](#)

简介

本文档介绍F2/F2e MAC表的完整条件和缓解该条件的方法。

每个芯片上交换机(SoC)的F2模块MAC限制为16k，报告随机MAC表包含60%利用率的完全错误消息。为什么线卡无法利用可用的整个16k MAC表空间？

```
%L2MCAST-SLOT2-2-L2MCAST MAC_FULLL LC: Failed to insert entry in MAC table for FE 1 swidx 271 (0x10f) with err (mac table full). To avoid possible multicast traffic loss, disable OMF. Use the con figuration CLI: "no ip igmp snooping optimise-multicast-flood"
```

先决条件

要求

思科建议您了解Nexus 7000架构。

使用的组件

本文档中的信息基于以下软件和硬件版本：

- Nexus 7000，带6.2.10及更高版本。
- F2e系列线卡。

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。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

背景信息

F2模块每个转发引擎SoC有16k MAC表空间。

每个模块上有12个此类SoC，每个服务有4个端口。

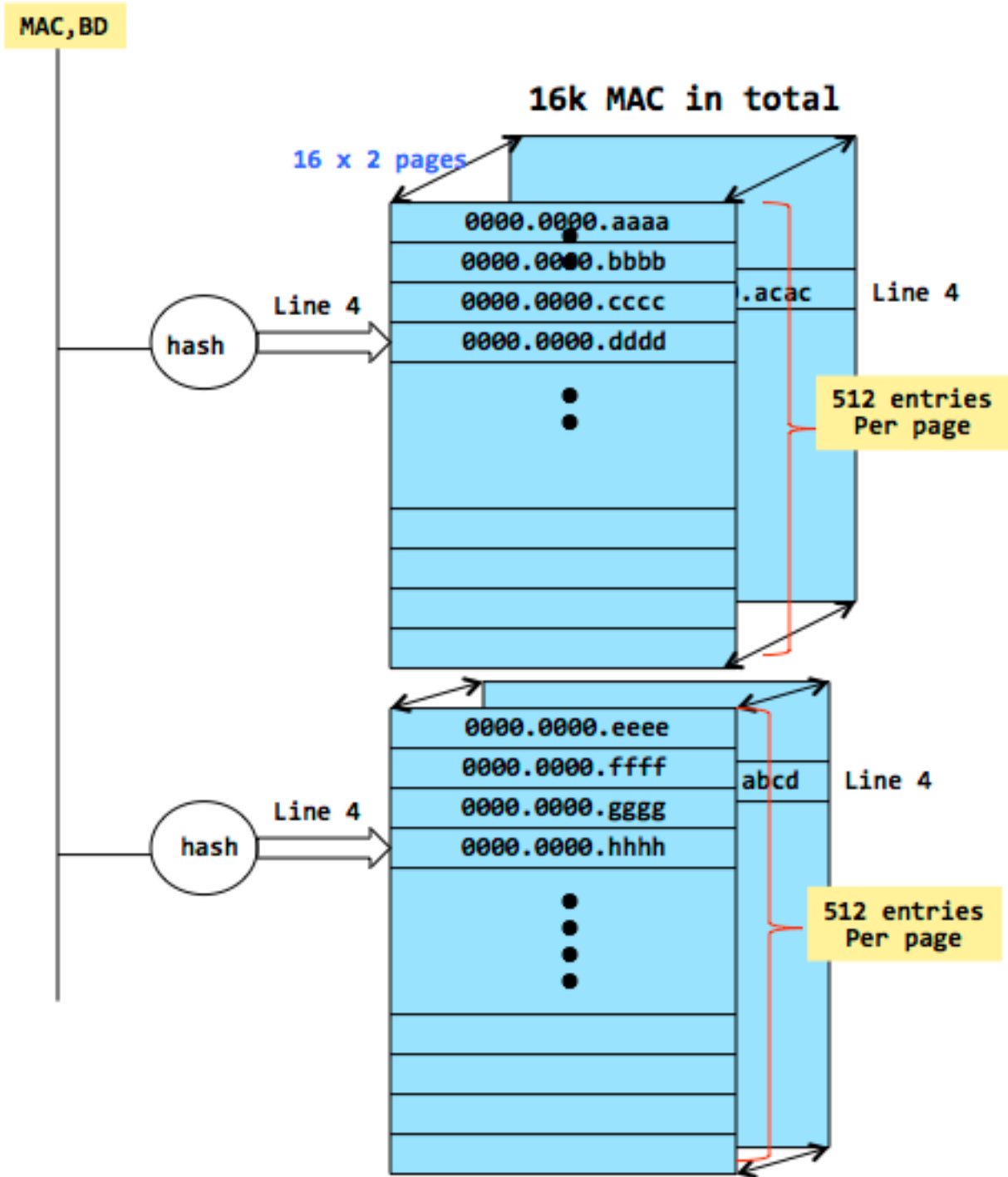
```
module-1# show hardware internal forwarding f2 l2 table utilization instance all
L2 Forwarding Resources
```

```
-----
L2 entries: Module inst  total    used  mcast  ucast  lines  lines_full
-----
```

L2 entries:	Module	inst	total	used	mcast	ucast	lines	lines_full
	1	0	16384	9647	265	9382	512	0
	1	1	16384	7430	1	7429	512	0
	1	2	16384	9654	264	9390	512	0
	1	3	16384	7430	7	7423	512	0
	1	4	16384	7564	8	7556	512	0
	1	5	16384	7432	1	7431	512	0
	1	6	16384	7418	0	7418	512	0
	1	7	16384	558	0	558	512	0
	1	8	16384	558	0	558	512	0
	1	9	16384	558	0	558	512	0
	1	10	16384	558	0	558	512	0
	1	11	16384	7416	0	7416	512	0

此处的输出重点显示了每个SoC硬件MAC地址表的使用情况。

要了解为什么会收到MAC表完整消息，您需要了解MAC表是如何划分的。此图可帮助您直观地了解情况。



- F2线卡的MAC表为16k，分页。每页可容纳512个条目。你一共有32页。您可以使用双向哈希将新MAC放入其中一个页面。
- 现在，我们假设每个页面都使用第4行。这意味着，32个唯一MAC最终产生了散列输出，将其置于每页的同一行。
- 如果生成具有相同哈希输出的33^{MAC}，则无法安装该MAC，并且可能会看到前面显示的错误消息。
- 行完整列跟踪已达到此状态的行数。

此处的输出还显示每页的行数，以及行满情况是否已达到。

```
module-2# show hardware internal forwarding f2 l2 table utilization instance all
```

L2 Forwarding Resources

L2 entries:							
Module	inst	total	used	mcast	ucast	lines	lines_full
2	0	16384	12280	283	11997	512	3
2	1	16384	12279	283	11996	512	2
2	2	16384	12289	283	12006	512	1
2	3	16384	12279	282	11997	512	2

只有最终以特定方式散列的MAC地址才会遇到此情况，而您不会看到其他MAC地址有任何问题。

通常，组播MAC地址可以更频繁地看到这一点，因为它们不像单播MAC那样随机。线卡通常使用行业标准RFC测试进行测试，以验证利用率。但是，特定客户环境中的某些MAC组合总有可能不能很好地优化，从而导致此错误。

缓解步骤

这些步骤有助于减少MAC表的使用。

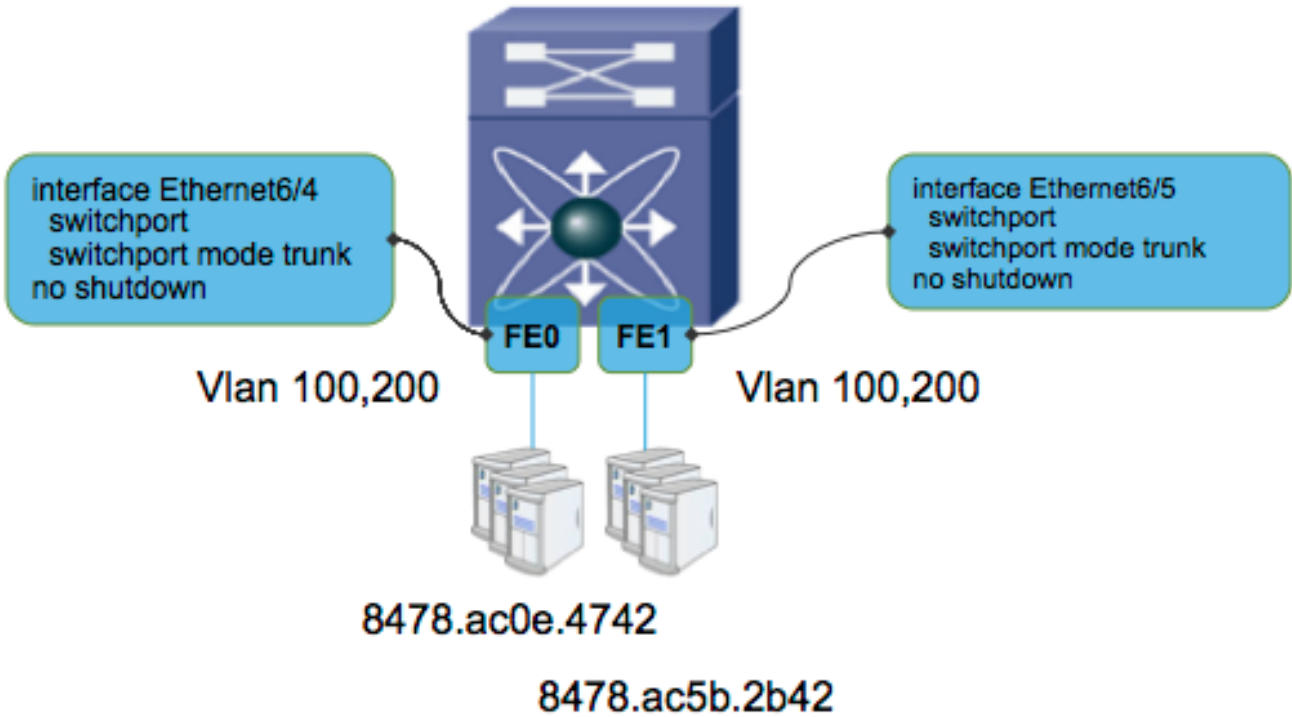
- 修剪VLAN
- L3分离
- 其他设计选项(fabricpath)
- M2或F3模块，用于未来增长

选项1.修剪VLAN

注意： VLAN 100和200没有SVI。这是一个重要假设，当您阅读选项2时，它将变得清晰。

在此简化设置中，不同SoC上有两台主机。

F2/F2e



```
N7KA-VDC-1(config-vlan)# sh mac address-table
```

Note: MAC table entries displayed are getting read from software.
Use the 'hardware-age' keyword to get information related to 'Age'

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen,+ - primary entry using vPC Peer-Link,
(T) - True, (F) - False , ~~~ - use 'hardware-age' keyword to retrieve age info

VLAN	MAC Address	Type	age	Secure	NTFY	Ports/SWID.SSID.LID
* 100	8478.ac0e.4742	dynamic	~~~	F	F	Eth6/4
* 200	8478.ac5b.2b42	dynamic	~~~	F	F	Eth6/5

```
N7KA-VDC-1# sh vlan internal bd-info vlan-to-bd 100
```

VDC Id	Vlan Id	BD Id
1	100	38

```
N7KA-VDC-1# sh vlan internal bd-info vlan-to-bd 200
```

VDC Id	Vlan Id	BD Id
1	200	39

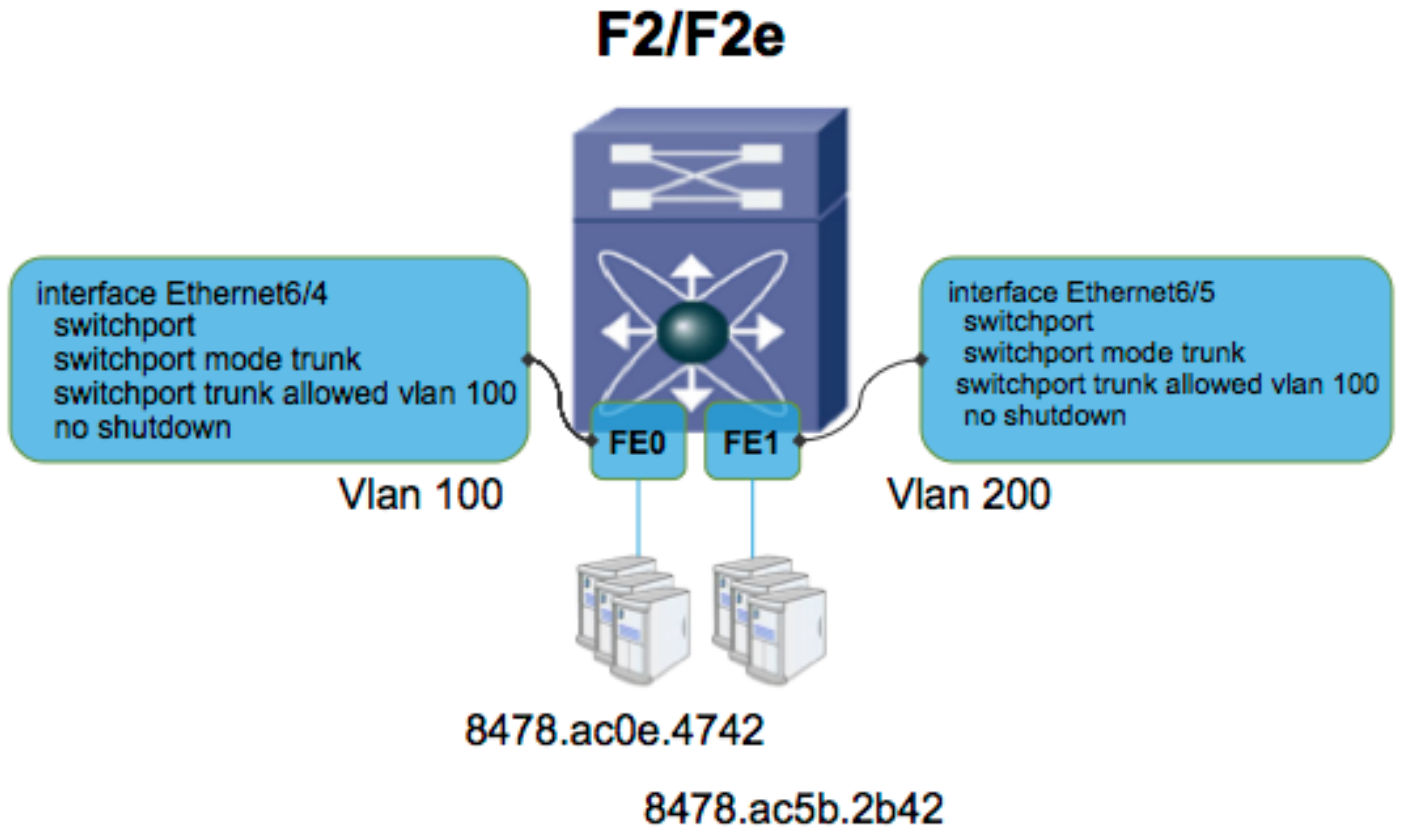
```
N7KA-VDC-1(config-if-range)# sh hard mac address-table 6
```

```
FE | Valid| PI| BD | | MAC | | Index| Stat| SW | Modi| Age| Tmr|
```

						ic		fied	Byte	Se1
0	1	1	38	8478.ac0e.4742	0x00053	0	0x081	1	138	1
0	1	0	39	8478.ac5b.2b42	0x00054	0	0x091	1	138	1
1	1	0	38	8478.ac0e.4742	0x00053	0	0x091	1	138	1
1	1	1	39	8478.ac5b.2b42	0x00054	0	0x081	1	138	1

每个FE (转发引擎= SoC) 显示2个MAC地址正在使用。

现在，您将修剪vlan，配置如下图所示。



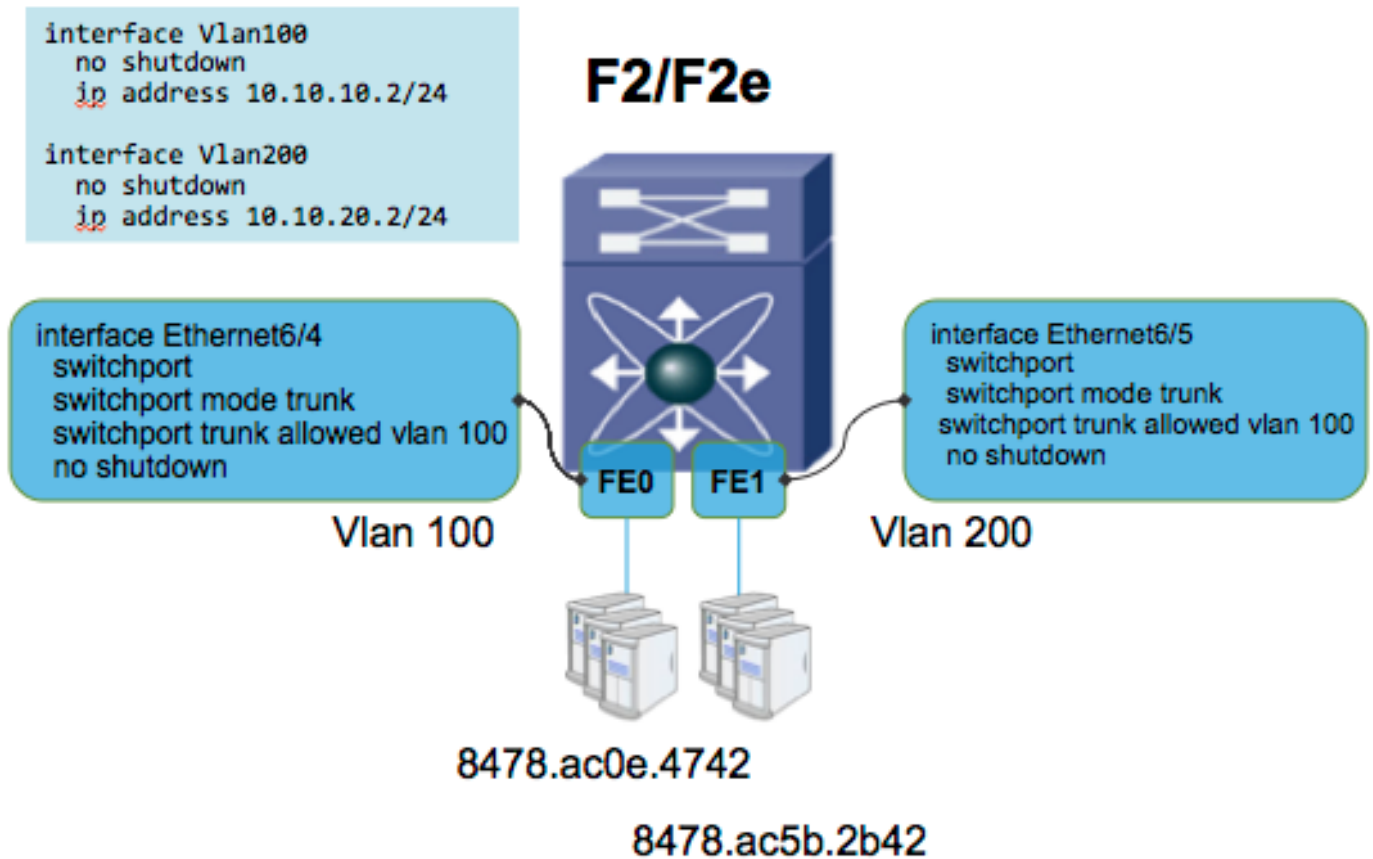
修剪VLAN后，每个FE(SoC)的条目数将减少一个。修剪VLAN会阻止MAC地址的FE之间同步。

```
N7KA-VDC-1(config-if-range)# sh hard mac address-table 6
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr
						ic		fied	Byte	Se1
0	1	1	38	8478.ac0e.4742	0x00053	0	0x081	1	138	1
1	1	1	39	8478.ac5b.2b42	0x00054	0	0x081	1	138	1

选项2. L3分离

在此，您已修剪了vlan，但假设您已在此VDC上为vlan 100和200配置了交换机虚拟接口(SVI)。



MAC表将如下所示，即使VLAN已修剪，MAC地址也会在FE之间同步。这是因为交换机虚拟接口(SVI)已启用，这要求FE也了解来自其他vlan的MAC地址。

```
N7KA-VDC-1(config-if-range)# sh hard mac address-table 6
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr
						ic		fied	Byte	Sel
0	1	1	38	8478.ac0e.4742	0x00053	0	0x081	1	138	1
0	1	0	39	8478.ac5b.2b42	0x00054	0	0x091	1	138	1
1	1	0	38	8478.ac0e.4742	0x00053	0	0x091	1	138	1
1	1	1	39	8478.ac5b.2b42	0x00054	0	0x081	1	138	1

如果删除VLAN 200 SVI，则MAC表不会看到FE0上VLAN 200 MAC的同步。

```
N7KA-VDC-1(config-if-range)# sh hard mac address-table 6
```

FE	Valid	PI	BD	MAC	Index	Stat	SW	Modi	Age	Tmr
						ic		fied	Byte	Sel
0	1	1	38	8478.ac0e.4742	0x00053	0	0x081	1	138	1
1	1	0	38	8478.ac0e.4742	0x00053	0	0x091	1	138	1

此步骤的结论不是删除SVI，而是分析是否可通过创建单独的第3层VDC将SVI移动到不同的VDC。这不是一个简单的设计步骤，需要详细规划。

选项3. Fabricpath等备用设计架构

这些是更复杂的备选方案，超出本文档的详细说明范围，但可以提高MAC使用效率。

选项4.使用M2/F3卡等高容量线卡

M2和F3线卡的MAC表容量要高得多。

[M2产品手册](#)==> MAC表 (每SoC 128k)

[F3产品手册](#)==> MAC表 (每SoC 64k)